

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

```
Run on:      April 5, 2004, 10:03:59 ; Search time 169 Seconds
              (without alignments)
              11947.271 Million cell updates/sec
```

Title: US-10-081-051-52

```
Perfect score: 4750
Sequence: 1 gatcctgctataacttag.....ccacaattgactaatgtatc 4750
```

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Database: N_Geneseq_29Jan04:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|-------------|--------------------|
| 1 | 4750 | 100.0 | 4750 | 6 | AAD88255 | Aad88255 Ehrlichia |
| 2 | 1416 | 29.8 | 1416 | 6 | AAD88257 | Aad88257 Ehrlichia |
| 3 | 597 | 12.6 | 597 | 6 | AAD88258 | Aad88258 Ehrlichia |
| 4 | 371 | 7.8 | 371 | 6 | AAD88256 | Aad88256 Ehrlichia |
| 5 | 260.4 | 5.5 | 1370 | 6 | ABK97886 | ABK97886 DNA encod |
| 6 | 187.4 | 3.9 | 1017 | 7 | ACA88500 | ACA88500 Prokaryot |
| 7 | 186.4 | 3.9 | 8056 | 7 | ABZ10246 | ABZ10246 Haematopo |
| 8 | 175.2 | 3.7 | 1029 | 7 | ACA29483 | ACA29483 Prokaryot |
| 9 | 170.6 | 3.6 | 996 | 7 | ACA47630 | ACA47630 Prokaryot |
| 10 | 168.6 | 3.5 | 996 | 7 | ACF57135 | ACF57135 Staphyloc |
| 11 | 168.6 | 3.5 | 999 | 7 | AA555337 | AA555337 Staphyloc |
| 12 | 168.2 | 3.5 | 999 | 7 | ACA27525 | ACA27525 Prokaryot |
| 13 | 165.8 | 3.5 | 110000 | 00 | AAK02428 | AAK02428 Borrelia |
| 14 | 165.8 | 3.5 | 110000 | 2 | AAK28248_01 | Continuation (2 of |
| 15 | 165.4 | 3.5 | 999 | 7 | ACA46288 | ACA46288 Prokaryot |
| 16 | 162.2 | 3.4 | 8056 | 7 | ABZ10246 | ABZ10246 Haematopo |
| 17 | 160 | 3.4 | 1557 | 8 | AD311919 | AD311919 DNA encod |
| 18 | 159.6 | 3.4 | 83391 | 6 | ABQ67093 | ABQ67093 Human an |
| 19 | 155.2 | 3.3 | 7814 | 4 | AA546530 | AA546530 Tumour su |
| 20 | 155 | 3.3 | 1065 | 7 | ACN33900 | ACN33900 Prokaryot |
| 21 | 155 | 3.3 | 1071 | 9 | ADC90723 | ADC90723 E. faeciu |
| 22 | 155 | 3.3 | 5641 | 6 | ABL33396 | ABL33396 Human imm |
| 23 | 154.2 | 3.2 | 918 | 2 | AAV74824 | AAV74824 Staphyloc |

| | | | | | | | |
|---|----|-------|-----|----------|---|-------------|---------------------|
| C | 24 | 153 | 3.2 | 15548 | 5 | AB1334155 | Ab1334155 Human imm |
| C | 25 | 150.6 | 3.2 | 8056 | 7 | ABE10100 | ABE10100 Haematop |
| C | 26 | 148.8 | 3.1 | AB154314 | 6 | AB154314 | AB154314 Chemical |
| C | 27 | 148.4 | 3.1 | 1020 | 7 | ACA18468 | ACA18468 Prokaryot |
| C | 28 | 148.4 | 3.1 | 1023 | 4 | AA53118 | AA53118 Enterococc |
| C | 29 | 147.2 | 3.1 | 1026 | 4 | AA55187 | AA55187 Enterococc |
| C | 30 | 147.2 | 3.1 | 6132 | 6 | AB132863 | AB132863 Human imm |
| C | 31 | 146.2 | 3.1 | 110000 | 6 | ABO69345_20 | Continuation (21 o |
| C | 32 | 146.2 | 3.1 | 319630 | 6 | ABO67194 | ABO67194 listeria |
| C | 33 | 145.4 | 3.1 | 11745 | 6 | ABE28332 | ABE28332 DNA trans |
| C | 34 | 144.2 | 3.0 | 7597 | 6 | AB133013 | AB133013 Human imm |
| C | 35 | 143 | 3.0 | 47108 | 6 | ABK1511 | ABK1511 Signal tr |
| C | 36 | 141 | 3.0 | 5611 | 6 | ABO67070 | ABO67070 Human ang |
| C | 37 | 139.8 | 2.9 | 1062 | 9 | ADG91772 | ADG91772 E. faecit |
| C | 38 | 139.8 | 2.9 | 2791 | 6 | ABO70865 | ABO70865 listeria |
| C | 39 | 139 | 2.9 | 61735 | 6 | AB133307 | AB133307 Human imm |
| C | 40 | 138.8 | 2.9 | 37515 | 6 | ABO6698 | ABO6698 Human ang |
| C | 41 | 137 | 2.9 | 18154 | 6 | AB133224 | AB133224 Human imm |
| C | 42 | 136.8 | 2.9 | 1008 | 4 | AA553251 | AA553251 Haemophil |
| C | 43 | 136.8 | 2.9 | 1008 | 7 | ACA34182 | ACA34182 Prokaryot |
| C | 44 | 136.8 | 2.9 | 110000 | 2 | AA424053_06 | Continuation (7 of |
| C | 45 | 136.2 | 2.9 | 11691 | 6 | AB134240 | AB134240 Human imm |

ALIGNMENTS

RESULT 1
 ID AAD48255 standard; DNA; 4750 BP.
 AC AAD48255;
 DT 24-FEB-2003 (first entry)
 DE Ehrlichia ruminantium 19hwoF2 protein encoding DNA.
 XX Vaccine; immunity; rickettsial infection; spotted fever; heart water;
 KW typhus; pathogen; immunostimulant; glycerol 3-phosphate dehydrogenase;
 KM 19hwoF2; antibacterial; enzyme; gene; ds.
 XX Ehrlichia ruminantium.
 OS
 Key Location/Qualifiers
 CDS 632..1615
 FT /*tag= a
 FT /product= "Hypothetical glycerol 3-phosphate
 FT dehydrogenase (19hwoF2)"
 FT /note= "This region is specifically claimed as SEQ ID NO:
 FT 54 in claim 1 of the specification"
 XX
 XX WO200266652-A2.
 PN
 PD 29-AUG-2002.
 PP 20-FEB-2002; 2002MC-US005772.
 XX
 XX 20-FEB-2001; 2001US-0269944P.
 PR
 XX
 XX (UYFL) UNIV FLORIDA.
 PA
 XX
 PI Barbet AF, Whitmore WM, Kamper SM, Simbi BH, Ganta RR;
 PI Moreland AL, Mwangi DM, Mwangi TC, Mahan SM;
 DR WPI; 2002-723186/78.
 DR P-PSDB; AAE31068.
 XX
 XX New Ehrlichia ruminantium polynucleotides, useful as vaccines for
 PT inducing protective immunity, and protecting animals or humans against
 PT rickettsial diseases, e.g. typhus, spotted fever or heart water.
 XX
 XX Claim 1; Page 131-133; 206pp; English.

CC The present invention relates to nucleic acid vaccines for conferring
CC immunity to rickettsial infection, including Ehrlichia ruminantium
CC (formerly Cowdria ruminantium). The invention also relates to novel E.
CC ruminantium polynucleotides and their corresponding proteins. Sequences
CC of the invention are useful for inducing immunity, particularly
CC protective immunity. They are also useful for detecting the presence of
CC E. ruminantium in a biological sample. They are useful in vaccines for
CC protecting animals or humans against rickettsial diseases, e.g. typhus,
CC spotted fever or heat water. Sequences of the invention are useful for
CC detecting antibodies to pathogens. The present sequence is a DNA encoding
CC Ehrlichia ruminantium hypothetical glycerol 3-phosphate dehydrogenase
CC (19kDorf2) protein

XX Sequence 4750 BP; 1835 A; 876 C; 496 G; 1543 T; 0 U; 0 Other;

Query Match 100.0%; Score 4750; DB 6; Length 4750;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4750; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATCTGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
DB 1 GATCTGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
QY 61 GCTGCTGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
DB 61 GCTGCTGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
QY 121 AGTACCTTGAATGCTGCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
DB 121 AGTACCTTGAATGCTGCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
QY 181 CTCGAATCTGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
DB 181 CTCGAATCTGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
QY 241 TTATGCAAGTCTGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
DB 241 TTATGCAAGTCTGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
QY 301 GAATGCTGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
DB 301 GAATGCTGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
QY 361 TTATGCAATCTGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
DB 361 TTATGCAATCTGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
QY 421 AAAAATCTGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
DB 421 AAAAATCTGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
QY 481 TATATTTGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
DB 481 TATATTTGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
QY 541 TATATTTGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
DB 541 TATATTTGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
QY 601 GATGCTGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
DB 601 GATGCTGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
QY 661 ATTGCTGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
DB 661 ATTGCTGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
QY 721 ACGTATCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
DB 721 ACGTATCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
QY 781 CACATATCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
DB 781 CACATATCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA

DB 781 CACATATCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
QY 841 CAACATATCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
DB 841 CAACATATCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
QY 901 AATACATCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
DB 901 AATACATCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
QY 961 CGAATATCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
DB 961 CGAATATCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
QY 1021 AATTTATCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
DB 1021 AATTTATCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
QY 1081 TATATCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
DB 1081 TATATCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
QY 1141 TATATCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
DB 1141 TATATCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
QY 1201 AATGATCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
DB 1201 AATGATCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
QY 1261 TATGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
DB 1261 TATGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
QY 1321 TATGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
DB 1321 TATGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
QY 1381 TATGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
DB 1381 TATGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
QY 1441 TATGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
DB 1441 TATGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
QY 1501 TATGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
DB 1501 TATGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
QY 1561 TATGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
DB 1561 TATGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
QY 1621 TATGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
DB 1621 TATGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
QY 1681 TATGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
DB 1681 TATGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
QY 1741 TATGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
DB 1741 TATGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
QY 1801 TATGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
DB 1801 TATGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
QY 1861 TATGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
DB 1861 TATGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
QY 1921 TATGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA
DB 1921 TATGCTAATACCTTAAAGTATGTTTAAAGTGTGCTTAAATTAATAAACA

QY 1921 CATATTTTACATAAATACTAATAAAGAACCAATACATATATTAATTAATTAAC 1980
DB 1921 CATATTTTACATAAATACTAATAAAGAACCAATACATATATTAATTAATTAAC 1980
QY 1981 ACCATTTTACATAGTAATATTAATAGATGCTGACATATTAATTAATTAAC 2040
DB 1981 ACCATTTTACATAGTAATATTAATAGATGCTGACATATTAATTAATTAAC 2040
QY 2041 CATATTTTACATAAATACTAATAAAGAACCAATACATATATTAATTAATTAAC 2100
DB 2041 CATATTTTACATAAATACTAATAAAGAACCAATACATATATTAATTAATTAAC 2100
QY 2101 TTGCTAATTAACCAAAAGGTTAAATTAACATCTCTGACAAATATGAAGATACAA 2160
DB 2101 TTGCTAATTAACCAAAAGGTTAAATTAACATCTCTGACAAATATGAAGATACAA 2160
QY 2161 TACAGCTCAAAATTAATTAACAAATTTCTAATTAACCAATCTCTAATTAAGAAC 2220
DB 2161 TACAGCTCAAAATTAATTAACAAATTTCTAATTAACCAATCTCTAATTAAGAAC 2220
QY 2221 AATAGTAATTAACATAGTAATATTAATAGATGCTGACATATTAATTAAGAAC 2280
DB 2221 AATAGTAATTAACATAGTAATATTAATAGATGCTGACATATTAATTAAGAAC 2280
QY 2281 TCGTATATTAATTAATTAACATTAACAAATTTTCAATCTCAACAAATCTCAAG 2340
DB 2281 TCGTATATTAATTAATTAACATTAACAAATTTTCAATCTCAACAAATCTCAAG 2340
QY 2341 TATTAATTAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2400
DB 2341 TATTAATTAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2400
QY 2401 ATATCATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2460
DB 2401 ATATCATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2460
QY 2461 CATATTTTACATAAATACTAATAAAGAACCAATACATATATTAATTAATTAAC 2520
DB 2461 CATATTTTACATAAATACTAATAAAGAACCAATACATATATTAATTAATTAAC 2520
QY 2521 AGCTTCATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTA 2580
DB 2521 AGCTTCATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTA 2580
QY 2581 AGCTTCATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTA 2640
DB 2581 AGCTTCATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTA 2640
QY 2641 TTTAGCACTTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTA 2700
DB 2641 TTTAGCACTTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTA 2700
QY 2701 TGGCTGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2760
DB 2701 TGGCTGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2760
QY 2761 TTGCTTAATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTA 2820
DB 2761 TTGCTTAATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTA 2820
QY 2821 TATATGATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTA 2880
DB 2821 TATATGATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTA 2880
QY 2881 CGGTATATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTA 2940
DB 2881 CGGTATATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTA 2940
QY 2941 ATATATTAATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTA 3000
DB 2941 ATATATTAATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTA 3000

QY 3001 AGACATATTAATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTA 3060
DB 3001 AGACATATTAATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTA 3060
QY 3061 TCCCTCTGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3120
DB 3061 TCCCTCTGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3120
QY 3121 ATATATTAATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTA 3180
DB 3121 ATATATTAATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTA 3180
QY 3181 ATATATTAATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTA 3240
DB 3181 ATATATTAATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTA 3240
QY 3241 TGGTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3300
DB 3241 TGGTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3300
QY 3301 TATTAATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTAATTA 3360
DB 3301 TATTAATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTAATTA 3360
QY 3361 CGATATTAATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTA 3420
DB 3361 CGATATTAATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTA 3420
QY 3421 GATATTTTACATAAATACTAATAAAGAACCAATACATATATTAATTAATTAATTAAC 3480
DB 3421 GATATTTTACATAAATACTAATAAAGAACCAATACATATATTAATTAATTAATTAAC 3480
QY 3481 AGATCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3540
DB 3481 AGATCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3540
QY 3541 ATATATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTAATTA 3600
DB 3541 ATATATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTAATTA 3600
QY 3601 ATATATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTAATTA 3660
DB 3601 ATATATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTAATTA 3660
QY 3661 ATATATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTAATTA 3720
DB 3661 ATATATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTAATTA 3720
QY 3721 AGCATATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTAATTA 3780
DB 3721 AGCATATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTAATTA 3780
QY 3781 TATATATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTAATTA 3840
DB 3781 TATATATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTAATTA 3840
QY 3841 AATATATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTAATTA 3900
DB 3841 AATATATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTAATTA 3900
QY 3901 TGAATATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTAATTA 3960
DB 3901 TGAATATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTAATTA 3960
QY 3961 CTATATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTAATTA 4020
DB 3961 CTATATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTAATTA 4020
QY 4021 TATATATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTAATTA 4080
DB 4021 TATATATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTAATTA 4080
QY 4081 AATATATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTAATTA 4140
DB 4081 AATATATTAACCAAAATTAACCAATTTGCTGATATTAATTAATTAATTAATTAATTA 4140

| | | |
|----|-----------------------------------------------------------------------------|---------------------------------------------------|
| XX | Barbet AP, Whitmire WM, Kamper SM, Simbi BH, Ganta RR; | |
| XX | Moreland AL, Mwambi DM, McGuire TC, Mahan SM; | |
| XX | WPI; 2002-723186/78. | |
| XX | New <i>Escherichia</i> ruminantium polynucleotides, useful as vaccines for | |
| XX | inducing protective immunity, and protecting animals or humans against | |
| XX | ricketsial diseases, e.g. typhus, spotted fever or heart water. | |
| XX | Claim 1, Page 137-138, 206pp; English. | |
| XX | The present invention relates to nucleic acid vaccines for conferring | |
| XX | immunity to rickettsial infection, including <i>Escherichia</i> ruminantium | |
| XX | (formerly Cowdria ruminantium). The invention also relates to novel E. | |
| XX | ruminantium polynucleotides and their corresponding proteins. Sequences | |
| XX | of the invention are useful for inducing immunity, particularly | |
| XX | protective immunity. They are also useful for detecting the presence of | |
| XX | E. ruminantium in a biological sample. They are useful in vaccines for | |
| XX | protecting animals or humans against rickettsial diseases, e.g. typhus, | |
| XX | spotted fever or heart water. Sequences of the invention are useful for | |
| XX | detecting antibodies to pathogens. The present sequence is a DNA encoding | |
| XX | E. ruminantium nitrogen assimilation regulatory protein (19nwort3) | |
| XX | Sequence 1416 BP; 472 A; 297 C; 161 G; 486 T; 0 U; 0 Other; | |
| XX | Query Match | 29.8%; Score 1416; DB 6; Length 1416; |
| XX | Best Local Similarity | 100.0%; Pred. No. 7.3e-215; Mismatches 0; Gaps 0; |
| XX | Matches 1416; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | |
| QY | 2467 TTATTTCAGAAACATTACCAATCCAGAAATTTCAATTTACGGTGAATGCTGAACGCTC | 2526 |
| DB | 1 TTATTTCAGAAACATTACCAATCCAGAAATTTCAATTTACGGTGAATGCTGAACGCTC | 60 |
| QY | 2527 CATCCCAAAATTCACAGAGTGGTGAATACCTACCTCCAAAAGAGATACGAGTTT | 2586 |
| DB | 61 CATCCCAAAATTCACAGAGTGGTGAATACCTACCTCCAAAAGAGATACGAGTTT | 120 |
| QY | 2587 TAAATACTGCTTTCAATTTCTCAGAGCTTTACGTAATGTAACAGAAATTAACCTTAGC | 2646 |
| DB | 121 TAAATACTGCTTTCAATTTCTCAGAGCTTTACGTAATGTAACAGAAATTAACCTTAGC | 180 |
| QY | 2647 ACTTAAACATCATTAATAGCGCAATTAAGTACTATATCTACGTAATCTTTTGGCT | 2706 |
| DB | 181 ACTTAAACATCATTAATAGCGCAATTAAGTACTATATCTACGTAATCTTTTGGCT | 240 |
| QY | 2707 AATCATCTCCTTAGAGAGATTCATATATTAATCAATTCATATACTTAACGTAATGCT | 2766 |
| DB | 241 AATCATCTCCTTAGAGAGATTCATATATTAATCAATTCATATACTTAACGTAATGCT | 300 |
| QY | 2767 TAAGTTCACCGGCATTCATATGACGTGATGCTATTAAGCTCATACTTAATACATG | 2826 |
| DB | 301 TAAGTTCACCGGCATTCATATGACGTGATGCTATTAAGCTCATACTTAATACATG | 360 |
| QY | 2827 AGTACACAAACCTAATTTTTTACAGATGCTATCATATAATACCTACCAATTCGGGAT | 2886 |
| DB | 361 AGTACACAAACCTAATTTTTTACAGATGCTATCATATAATACCTACCAATTCGGGAT | 420 |
| QY | 2887 ATCTGTACAAATATCTACTAAGACGGTACTCTAATGGAAGACATTAATCTATATA | 2946 |
| DB | 421 ATCTGTACAAATATCTACTAAGACGGTACTCTAATGGAAGACATTAATCTATATA | 480 |
| QY | 2947 TAAATCTCAGAAAACCTACAGGCTTTTACTTCACTTCAATATCTTTGGAAGAAGAC | 3006 |
| DB | 481 TAAATCTCAGAAAACCTACAGGCTTTTACTTCACTTCAATATCTTTGGAAGAAGAC | 540 |
| QY | 3007 AATATATTCACATCTATACATAAGAGAACTTACTATTTCCCTATATATTTTCCCTC | 3066 |
| DB | 541 AATATATTCACATCTATACATAAGAGAACTTACTATTTCCCTATATATTTTCCCTC | 600 |
| QY | 3067 CTGTATATATCTGAGTAATCTTAATGGCGTATGCTATGTAATCTGTACTCATCAT | 3126 |
| DB | 601 CTGTATATATCTGAGTAATCTTAATGGCGTATGCTATGTAATCTGTACTCATCAT | 660 |

QY 3127 AAAAAGGACGATGATTTGGTCTGCTATTAATCCAAATATGAGAGGATCTGATAGA 3186
 DB 661 AAAAAGGACGATGATTTGGTCTGCTATTAATCCAAATATGAGAGGATCTGATAGA 720
 QY 3187 CAATATATATATATCTTCTGCTACGACCAATATATTAACCAATATATATATCTGAT 3246
 DB 721 CAATATATATATATCTTCTGCTACGACCAATATATTAACCAATATATATATCTGAT 780
 QY 3247 CATGATGAGTACATATATATATATGAGATGATCAATCCCTGATTTTATATATAG 3306
 DB 781 CATGATGAGTACATATATATATATGAGATGATCAATCCCTGATTTTATATATAG 840
 QY 3307 CCTGCAACTACTCTCTTCCACACCTGGGCAACGATATAGATATGATGATGATGAT 3366
 DB 841 CCTGCAACTACTCTCTTCCACACCTGGGCAACGATATAGATATGATGATGATGAT 900
 QY 3367 AGTAGCTGCTTATTAATCACTACTCTCAAAATTAAGATATACAGGGAGTTACCGACTAT 3426
 DB 901 AGTAGCTGCTTATTAATCACTACTCTCAAAATTAAGATATACAGGGAGTTACCGACTAT 960
 QY 3427 TTCAATATCTCAATATGCTATTTCAATCACTATTTCTCTCACTATCTATCACTAC 3486
 DB 961 TTCAATATCTCAATATGCTATTTCAATCACTATTTCTCTCACTATCTATCACTAC 1020
 QY 3487 TATAGCTCTCTTACAACTACTTAACTCTCTCTGTAAGGCTTTTCTATATATATC 3546
 DB 1021 TATAGCTCTCTTACAACTACTTAACTCTCTCTGTAAGGCTTTTCTATATATATC 1080
 QY 3547 ATAAGCACCATATGACGAGACTTTAGACGAGTGGCAATTTTCCATGCCACTATATC 3606
 DB 1081 ATAAGCACCATATGACGAGACTTTAGACGAGTGGCAATTTTCCATGCCACTATATC 1140
 QY 3607 AATAAGCAGCAATATGAGATACCTTTCTTATAGCTTTTCAAGACCTTATCATCAT 3666
 DB 1141 AATAAGCAGCAATATGAGATACCTTTCTTATAGCTTTTCAAGACCTTATCATCAT 1200
 QY 3667 ATCAAGATCTCTTACCAATATATCAATATATATCAATCAATCAATCAATCAATCAAT 3726
 DB 1201 ATCAAGATCTCTTACCAATATATCAATATATATCAATCAATCAATCAATCAATCAAT 1260
 QY 3727 CTGATGCGGATTAACCAATCAATCTGATATTTAGTACATATATATATATATATATAT 3786
 DB 1261 CTGATGCGGATTAACCAATCAATCTGATATTTAGTACATATATATATATATATATAT 1320
 QY 3787 ATCTTTATATAGATTTCTGATATCAATCTGATATCAATCAATCAATCAATCAATCAAT 3846
 DB 1321 ATCTTTATATAGATTTCTGATATCAATCTGATATCAATCAATCAATCAATCAATCAAT 1380
 QY 3847 CAATCTTTCTTGGAGATTTCAAAATCTCTGCTCAT 3882
 DB 1381 CAATCTTTCTTGGAGATTTCAAAATCTCTGCTCAT 1416

RESULT 3
 AAD48258
 ID AAD48258 standard; DNA; 597 BP.
 AC AAD48258;
 DT 24-FEB-2003 (first entry)
 DE Ehrlichia ruminantium 19hwoF41 protein encoding DNA.
 XX
 KM Vaccine; immunity; rickettsial infection; spotted fever; heart water;
 KM typhus; pathogen; immunostimulant; antibacterial; 19hwoF41; gene;
 KM cell division protein; ftsQ; ds.
 XX Ehrlichia ruminantium.
 OS
 XX
 PN WO20026652-A2.
 XX
 PD 29-AUG-2002.

XX 20-FEB-2002; 2002WO-US005772.
 PF
 XX 20-FEB-2001; 2001US-0269944P.
 PR
 XX
 XX (UWFL) UNIV FLORIDA.
 PA
 XX Barbet AF, Whitmore WM, Kamper SM, Simbi BH, Ganta RR;
 PI Moreland AL, Wmangi DM, McGuire TC, Mahan SM;
 XX WPI; 2002-723186/78.
 DR
 XX
 PT New Ehrlichia ruminantium polynucleotides, useful as vaccines for
 PT inducing protective immunity, and protecting animals or humans against
 PT rickettsial diseases, e.g. typhus, spotted fever or heart water.
 XX
 PS Claim 1, Page 138-139; 206pp; English.
 CC The present invention relates to nucleic acid vaccines for conferring
 CC immunity to rickettsial infection, including Ehrlichia ruminantium
 CC (formerly Cowdria ruminantium). The invention also relates to novel E.
 CC ruminantium polynucleotides and their corresponding proteins. Sequences
 CC of the invention are useful for inducing immunity, particularly
 CC protective immunity. They are also useful for detecting the presence of
 CC E. ruminantium in a biological sample. They are useful in vaccines for
 CC protecting animals or humans against rickettsial diseases, e.g. typhus,
 CC spotted fever or heart water. Sequences of the invention are useful for
 CC detecting antibodies to pathogens. The present sequence is a DNA encoding
 CC E. ruminantium hypothetical cell division protein (ftsQ, 19hwoF41)
 XX
 SQ Sequence 597 BP; 220 A; 117 C; 68 G; 192 T; 0 U; 0 Other;
 Query Match 12.6%; Score 597; DB 6; Length 597;
 Best Local Similarity 100.0%; Pred. No. 1,9e-85;
 Matches 597; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4154 TTATTTCTATATATTAACCTTCTGGAACGCGATGCTATAGTTTCCATATCAAAA 4213
 DB 1 TTATTTCTATATATTAACCTTCTGGAACGCGATGCTATAGTTTCCATATCAAAA 60
 QY 4214 CTCACCTTAACCTTATATGATATATCAATATATTCAGCGACGACATATATATAT 4273
 DB 61 CTCACCTTAACCTTATATGATATATCAATATATTCAGCGACGACATATATATAT 120
 QY 4274 AGGTATATTTATATTTAAACCACTAGACAAATATATATCCACCAATGACTATCAACATA 4333
 DB 121 AGGTATATTTATATTTAAACCACTAGACAAATATATATCCACCAATGACTATCAACATA 180
 QY 4334 TGTGATGATGAACCAATACCAACCAACAGATATATATATATATATATATATATATAT 4393
 DB 181 TGTGATGATGAACCAATACCAACCAACAGATATATATATATATATATATATATATAT 240
 QY 4394 ATCTAGATGATTAACGACCATCTCCATGATTTGAAGTCAATCATCCGATATATTA 4453
 DB 241 ATCTAGATGATTAACGACCATCTCCATGATTTGAAGTCAATCATCCGATATATTA 300
 QY 4454 ATATATCAATATATATATGTCACAAACATATCAATGATATTTTGTATATATATATCA 4513
 DB 301 ATATATCAATATATATATGTCACAAACATATCAATGATATTTTGTATATATATATCA 360
 QY 4514 ATTTGCAAAAGCAGATATCTTGAACGTATATCTGTAAGATATTTGTAACAGCTCTT 4573
 DB 361 ATTTGCAAAAGCAGATATCTTGAACGTATATCTGTAAGATATTTGTAACAGCTCTT 420
 QY 4574 AACTGAGGATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4633
 DB 421 AACTGAGGATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 QY 4634 AGGCACAAAAAGATAGACCTTGATGATGATGATGATGATGATGATGATGATGATGAT 4693
 DB 481 AGGCACAAAAAGATAGACCTTGATGATGATGATGATGATGATGATGATGATGATGAT 540
 QY 4694 TTCAATACCATCAATCAATATTTTATCAACAGTAAATCCAAATGACTAATGATC 4750

Db 541 TTGATTACGATCAATTTTATTCACAGTAATTCACAAATTGACTAATTGATC 597

RESULT 4

AAD48256 standard; DNA; 371 BP.

XX AAD48256;

DT 24-FEB-2003 (first entry)

XX Ehrlichia ruminantium 19hwoxfli protein encoding DNA.

XX Vaccine; immunity; rickettsial infection; spotted fever; heart water;

XX typhus; pathogen; immunostimulant; integrase; recombinase; 19hwoxfli;

XX antibacterial; enzyme; gene; ds.

XX Ehrlichia ruminantium.

XX WO200266652-A2.

XX 29-AUG-2002.

XX 20-FEB-2002; 2002WO-US005772.

XX 20-FEB-2001; 2001US-0269944P.

XX (UYFL) UNIV FLORIDA.

XX Barbet AF, Whitmore MW, Kamper SM, Simbi BH, Ganta RR,

XX Moreland AL, Mwangi DM, McGuire TC, Mahan SM;

XX WPI; 2002-723186/78.

XX New Ehrlichia ruminantium polynucleotides, useful as vaccines for

XX inducing protective immunity, and protecting animals or humans against

XX rickettsial diseases, e.g. typhus, spotted fever or heart water.

XX Claim 1; Page 134; 206pp; English.

XX The present invention relates to nucleic acid vaccines for conferring

XX immunity to rickettsial infection, including Ehrlichia ruminantium

XX (formerly Cowdria ruminantium). The invention also relates to novel E.

XX ruminantium polynucleotides and their corresponding proteins. Sequences

XX of the invention are useful for inducing immunity, particularly

XX E. ruminantium in a biological sample. They are useful in vaccines for

XX protecting animals or humans against rickettsial diseases, e.g. typhus,

XX spotted fever or heart water. Sequences of the invention are useful for

XX detecting antibodies to pathogens. The present sequence is a DNA encoding

XX protein

XX Sequence 371 BP; 128 A; 74 C; 45 G; 124 T; 0 U; 0 Other;

XX

XX

XX

XX

XX

XX

XX

XX

XX

Db 181 CTCGAATCTCCAAATTTTAAACGACAAAGATTATTATATGTAACGAAACCCGTACTC 240

XX 241 TTATGCAAGAACTCTATPAACTTATCAAGGCTCTTACATAGAAACCACTGATTCGAA 300

XX 241 TTATGCAAGAACTCTATPAACTTATCAAGGCTCTTACATAGAAACCACTGATTCGAA 300

XX 301 GAATACCTCTTCTCTTGTTCATCCATATTTTCCAAATTTATTTACATATATGAGTCT 360

XX 301 GAATACCTCTTCTCTTGTTCATCCATATTTTCCAAATTTATTTACATATATGAGTCT 360

XX 361 TTATTCGACAT 371

XX 361 TTATTCGACAT 371

XX ABR97886 standard; DNA; 1370 BP.

XX ABR97886;

XX 10-OCT-2002 (first entry)

XX DNA encoding Lawsonia intracellularis antigenic ntrC protein.

XX Immunogen; fibs; filr; ntrC; glnH; motB; clyC; ytfM; ytfN; porcine;

XX pig; avian; bird; porcine proliferative enteropathy; PPE;

XX intestinal adenomatosis complex; porcine intestinal adenomatosis; PIA;

XX necrotic enteritis; proliferative haemorrhagic enteropathy;

XX regional ileitis; haemorrhagic bowel syndrome; vaccine; antibacterial;

XX porcine proliferative enteritis; Campylobacter spp.-induced enteritis;

XX gene; ds.

XX Lawsonia intracellularis.

XX WO200238594-A1.

XX 16-MAY-2002.

XX 09-NOV-2001; 2001WO-AU001462.

XX 10-NOV-2000; 2000AU-00001381.

XX 17-NOV-2000; 2000US-0249596P.

XX (AGRI-) AGRIC VICTORIA SERVICES PTY LTD.

XX (AUPO-) AUSTRALIAN PORK LTD.

XX (PFI2) PFIZER PROD INC.

XX Rosey EL, King KW, Good RT, Strugnell RA;

XX WPI; 2002-557448/59.

XX P-PSDB; ABG68912.

XX New immunogenic polypeptide comprising epitope of Lawsonia spp.

XX polypeptide such as fibh, filr, ntrC, glnH, motB, polypeptides, useful in

XX vaccines for treatment of porcine proliferative enteropathy in pigs and

XX birds.

XX Claim 6; Page 104-107; 155pp; English.

XX The invention describes an isolated or recombinant immunogenic

XX polypeptide (I) which comprises, mimics or cross-reacts with a B-cell or

XX T-cell epitope of a Lawsonia spp. polypeptide such as fibh, filr, ntrC,

XX glnH, motB, clyC, ytfM or ytfN polypeptides. (I) is useful for

XX identifying whether or not a porcine or avian animal has suffered from a

XX past infection, or is currently infected, with Lawsonia spp. or a

XX microorganism that is immunologically cross-reactive with Lawsonia spp.

XX Antibodies are useful for diagnosing infection of a porcine or avian

XX animal by Lawsonia spp. or a microorganism that is immunologically cross-

XX reactive with Lawsonia spp. A nucleic acid encoding a Lawsonia spp.

XX immunogen is useful as probes or primers for detecting Lawsonia spp. or

XX related microorganism in a biological sample derived from a porcine or

avian animal subject. (1) is preferably useful for vaccinating porcine animals against intestinal diseases collectively known as porcine proliferative enteropathy (PPE), previously known as intestinal adenomatosis complex, porcine intestinal adenomatosis (PIA), necrotic enteritis, proliferative haemorrhagic enteropathy, regional ileitis, Campylobacter spp.-induced enteritis. (1) is also useful in vaccines for the prophylaxis and treatment of PPE in birds. This sequence encodes a Lawsonia intracellularis immunogenic peptide used in the creation of a porcine vaccine described in the invention

Sequence 1370 BP; 443 A; 201 C; 278 G; 448 T; 0 U; 0 Other;

Query Match 5.5%; Score 260.4; DB 6; Length 1370;
Best Local Similarity 52.1%; Pred. No. 3.2e-32;
Matches 708; Conservative 0; Mismatches 631; Indels 19; Gaps 5;

2482 ACAACATCCAGATTTTCATTTAGCGTGAATGCTGAACGTTCCATTCCACCAATTC 2541
1358 ATAAATACATAGCTTTTAACTTTCTATATAAATTAATCTTTCAAGTCAATAGCTTC 1299
2542 AGCAGTTCGTATACATTTACCTCCAAACGAGATTAAGTATTAAATAGTGTCTTC 2601
1298 TGTATATGGGTATATTTCTTGTATAGCATGTAATTTTCAGTTAAATTTGGTTC 1239
2602 AAATTTCCAGACCTTACGTAATGCTA---GAATTAACCTTACGACTTAAACATC 2658
1238 AAAAGCTATTTTACCGGTAAATCTATATCGGGAAGAAAGTTTAAAGGAAAGAAA 1179
2659 ATTATAGCGCATTAAGTATATCTATCTAGTAAATCTTTGGTGAATCATCTCCTT 2718
1178 TTCTTGGTAAATTTTAAATTTGCTTTTACTTATGATCTGTCAACATCTTCTT 1119
2719 AGGAGATTTCAATTAATTAATCCATTTATACATTAAGTATGTTTAAAGTACCTG 2778
1118 CCGTGAATTAAGATTAACCATCTGCTTCAAAATTAAGTAAATCTTCAATCTCCCTG 1059
2779 CCATTATAGCATGCTATGCTATTTAAAGTTCATCACTTAATACATGATACAAACC 2838
1058 CCAACATATTTGTTCAATACAGGAGACCTCATTAATAAATTAAGCGCTCACGCT 999
2839 TATTTTTCAGATGCTATTCATTAATACTACACATTCGGATATCTGTACATA 2898
998 ATAACTTTACTCAATGATCAACCAACTATTTAATTAATAGTCAATACAGATCAG 939
2899 TTCTACAAAGAGGCTCTAATTTGGAAGACATTTAATCTATATTAATCTCCACA 2958
938 TTACGAGAGGGGAGATGATGACATGACATCGTAAGCGTAATACAAATCTTGAAG 879
2959 AAACCTACAGCTTTTACTTCACTTCAATATCTTTGGAAGAACACATTAATCTCAC 3018
878 AAATGTTCCATCGCTATAGCGCTTCAAGATCTTATTTGCTGCAATACCTTTC 819
3019 ATCTATCTACAGGAATCTTACTATTTTCCCTATATATTTTCCCTCTGTATATCT 3078
818 ATCAACTTTATAGTTTCAACCTACCAATTTTTCAAAACATTTCTTGGCAATAG 759
3079 GAGTAATCTTAATGCTATGCTATGTAATCGTAACTGTAATCTAATTAATTAAGCTAC 3138
758 CAATATTTTGTGCTTTTAAATCTCATATCTCTATTTCAAGAAATTAATGTTCC 699
3139 ATGATTTGCTGCTCTATTAATTCATATGAGAGGCTCTATGACATATATTAAT 3198
698 TTATATGCGCACTCAAAAGACCTGACAGAGATGCGCACGATTAAGGCC----- 644
3199 ACTTTCCTACATCAAAATATTAACCAAGTATTAATAGCTGATAGATGATAGTA 3258
643 -CCTTTTTCATGACCAATATGTTGCTTCAATCAATCTTCAGGAGATGACGACAAAT 585
3259 CATGATATAATGAGATCATACCCCTTGATTTTATGATTAAGCTTACCAATC 3318
584 AAGAGCTATAAATGTTTGTATATGATGAGCTTCTTGTGCAATGCTTGTGAGCTTA 525

3319 TTCCTTCCACACCTGGGAAACGATATGATATACGATCTGATGATAGCTTCTT 3378
524 CTTTACCTGTACATTTCTCTCTGTAAGTATGCCAAGCATCTGTGAGCACTG 465
3379 ATTAATCATCTTCTCAATTAAGTATTAACAGGAGGTACCACTATTAATCTCTC 3438
464 TATATATACTTTAAATTTTAATTAACAGGAGCTGTCTAATACTCACTCTCTC 405
3439 A---AATGCTATTTCACTATCATTTTCTCTATGATATCTACAGACTTATAGCTCT 3495
404 AGTATATACAGTACGATTAATTTGTTTCCCTTTAATCTTACTGTTTCTATAGCTCT 345
3496 CTTTACAACTAATCTTATCTCTCTGTAAGGCTTTCTTATATATCATAGACAC 3555
344 ATTACCTGTAAATTAAGACTTTTCCAAAGAAAGGCTTTTCAATTAATCATAGCAC 285
3556 CATATGACAGACTTTTAACAGATGCGCAATTAACATGCCACTAATCATATACAG 3615
284 TTGAGGATAGCTGTATACAGCATTTCAATGTGCGATGACCTGAATCATATACAG 225
3616 CAATTAAGATACCTTTCTTAAAGCTTTTCCAGTACACTTATCATCATATATGATCC 3675
224 TAACTCTGATGAGAGCTGATATGCTTAAAGCCATTAAGACCATCATCCAG---- 169
3676 TCTTAACCATATATGCAATTAATTAACATCAAGCTTTTTCATTAAGCAATCTGATCC 3735
168 --GAAGCCAAATATCAAGAAACAGATCTGAGACTTACATCAACACACTTAAGTCC 111
3736 GGATTAACCATCACTGCTAATTTAGTACATATTAATCATCACTAATATATCTTAT 3795
110 TTCTTCTGCTAAGCTCTTTCTAATTAATCTTATGAGCCCTCATCTTCAAAATTCCTTTCA 51
3786 TAGATTTCTGATATCACTTATCATCATCAATCAACACTAATA 3833
50 T-GAATCTAATAGAGCTTTCATCATCTATATATAGTA 14

RESULT 6

ACA28500 standard; DNA; 1017 BP.

ACA28500;

19-JUN-2003 (first entry)

Prokaryotic essential gene #10157.

Antisense; ds; prokaryotic essential gene; cell proliferation;

drug design; gene.

Clostridium botulinum.

MO200277183-A2.

03-OCT-2002.

21-MAR-2002; 2002MO-US009107.

21-MAR-2001; 2001US-00815242.

06-SEP-2001; 2001US-00948993.

25-OCT-2001; 2001US-0342923P.

08-FEB-2002; 2002US-00072851.

06-MAR-2002; 2002US-0362699P.

(ELIT-) ELITRA PHARM INC.

Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GT, Yamamoto R, Forsyth RA, Xu HH; WPI; 2003-029926/02. P-PSDB; ABU24630. New antisense nucleic acids, useful for identifying proteins or screening

PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.

Claim 14; SEQ ID NO 16370; 1766bp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pcc_sequences

XX Sequence 1017 BP; 412 A; 106 C; 229 G; 270 T; 0 U; 0 Other;

XX Query Match 3.9%; Score 187.6; DB 7; Length 1017;

XX Best Local Similarity 52.1%; Pred. No. 1e-20;

XX Matches 499; Conservative 0; Mismatches 444; Indels 15; Gaps 3;

QY 635 AAAATCAGTATTGTTGAGTGCAGATCATTGGCAGCAGCAATGCACTGTCAGCA 694
DB 31 AAGAGTTGTTTGTGAGCAGGTAGTTTGGACAGCTTTAGCTGATCTGAGAAA 90
QY 695 CATGGATATCAGTAACTTATGAGGAGCGTATGATGAAATATATACATATAAAGCT 754
DB 91 AAGAGTTGAAAGTTTACATATGCAATGCAATTAATACATATGAGAGATTAATATTC 150
QY 755 TACCGAAAAATTTAAATATTTTACCCACATATCATCTTACAGACACATATATGCAAC 814
DB 151 AAAAAAGAAAAATATAATTAATTTACCTAAGTAGTATACCTGAGGGGTAAAGCTTAT 210
QY 815 AGCAATATAGCAAGATGATTATCTGACACACATACATGATATATCTTACATATCTTCA 874
DB 211 AATGAAATGAAAAAGATTA---GACATATGTAATTTATATGTTAGCAGTACTCTCC 267
QY 875 CAACATTTAGCAGCATATGATGACAAATATACAAACAAACAGCATATGTGTAATAATCT 934
DB 268 CATGTGATAGACAGGTATGCAAAAAAGATAC-----AACCTTTATATAAGAAAAATCAA 321
QY 935 CCAATATTAATTTAGTAAAGGTATGCAAAATTAATCATCACTCAATTTCCAGAGGAATA 994
DB 322 ATATATGATAGTATGCTTAAAGGTATGAGAGGGGTGAGGAAAAAGTTATCAGAGGTC 381
QY 995 GGAGAGAAATTTTCAATATATATCCAAATTTTATCTCTGCTGCAAGTTTGTCTAAA 1054
DB 382 ATTAAGAAAGAACTGCTTAATATCTCTATAGTATGATTTGTCAGGCCCTTACAGCAGAG 441
QY 1055 GAAATTTGAGAACATCTTCTGTTGATATGATCTGCTGGATTAATAAAGAACTTGGT 1114

RESULT 7

ABZ10246/c

ID ABZ10246 standard; DNA; 8056 BP.

AC ABZ10246;

DT 16-JUN-2003 (first entry)

DE Haematopoietic cell proliferation disorder related DNA sequence #386.

XX Human; haematopoietic cell proliferation disorder; cytostatic;

KM gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;

XX cytosine methylation state; gene; ds.

OS Homo sapiens.

PN WO200277272-A2.

PD 03-OCT-2002.

PF 26-MAR-2002; 2002MO-BE003401.

PR 26-MAR-2001; 2001US-0278333P.

PA (EPIC-) EPIDENOMICS AG.

PI Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;

PI Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Liu E;

PI Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T, Pellet C;

PI Schwöpe I, Ziebarth H;

DR WPI; 2003-018942/01.

XX Detecting and differentiating between haematopoietic cell proliferative

PT disorders, comprises contacting a target nucleic acid with a reagent that

PT distinguishes between methylated and non-methylated CpG dinucleotides.

XX Claim 28; SEQ ID NO 386; 117bp; English.

DB 442 GAAAGTCTCAGATATACCTTACACAGTAGTGGCTTTCAGAGCATATGATATAGCT 501

QY 1115 GAATCATTTGATGAAACAAATTAAGTATGATGTTCTTAAATATATATACCATCAGATATTT 1174

DB 502 GAAAGTCTCAGATATACCTTATGATATATATTAATTTAGAGTATGATATGATGTTTA 561

QY 1175 ATGCTGTACAGATTGGAGCTGATTAAGAACATATATGCAATTCAGATGTTGAATATC 1234

DB 562 GTGAGGTGAAATAGAGAGCGCTGTAAATATATATGCTTTAGCTGAGAGGCTATCT 621

QY 1235 GCTGAAAAAATTTAGTATATGCTGTGCTACTGTTATTAATCAATTAAGCATGATGAA 1294

DB 622 GATGTATGATGTTATGATGATATACCAAGGACGCTTATGACAAAGATATGATGTA 681

QY 1295 ATTAAGACATATATATGCAAAAAATCATTTCAATGATCTTCATATTAATGTTGCA 1354

DB 682 ATATATAGA-----ATAGTCAAAAGCTTGAGAGTTAAAGAGAACTTTTCAGGCTCT 735

QY 1355 TCATGCTTGGAGATCTAATATTAATTAATGATGATGACAGCAAGCATTCAGCAATATGCTTT 1414

DB 736 ACAGGATGAGGATTTAATGATATGATGATGATGATGATGATGATGATGATGATGAT 795

QY 1415 GGACTGAAATAGGAAAGGTAGAAATATATATATATATATATATATATATATATAT 1474

DB 796 GGTATCTATATCGTTAAAGGCTATATGACAGAAAGCAATAGAAAGAGGTATGATGTA 855

QY 1475 GTTGAAGGAAACCGTACTGTTAAACCGATATATATATATATATATATATATATATAT 1534

DB 856 GTGAGGAAATTAAGCTTGTATATATATATATATATATATATATATATATATATATAT 915

QY 1535 CTACCAATTTGATATCTATTTTACATTTTATACATGATGATATATATCACTAGATTAAG 1592

DB 916 ATCCCTATTAACAGATCATTAATTAAGTTATTGAAAAATTAAGAGCCTATATATG 973

SQ Sequence 8056 BP; 3711 A; 0 C; 371 G; 3974 T; 0 U; 0 Other;

Matches 1900; Conservative 0; Mismatches 2341; Indels 99; Gaps 21;

| | | | |
|----|------|-----------------------------------------------------------------------|------|
| QY | 964 | AATTCAGACTCAAAATTTCCAGTGAATGTGCGAAGAAATTTTACATATATTCAT | 102 |
| DB | 3999 | TTTAACAATTTTCATTTTTTTTTTTTAATATACATAAATTTATATTTAAATTTTAT | 3940 |
| QY | 1024 | TTTATACCTCTGTCGCAAGTTTGTCTAAAGAAATGTGAGACATCTTCCTGTAGAT | 1083 |
| DB | 3939 | AAATTAATTTTTTATCATTTATTTTTCACAATTTTTTAAAA-----CATTAATAATCA | 3887 |
| QY | 1084 | AGTACTGCTGCGATATAAAGAACTGGAGATCATGTGATAGAAACAATAAGTAAATGA | 1143 |
| DB | 3886 | ATTTCOAATTTTATTCATTTCAATTCAAAATCATTTTCAATTTTACATTTTCAAAATTTTCAA | 3822 |
| QY | 1144 | TGTTCTAAAAATATATACATCAAGATATTTATAGTGTACAGATTTGAGCTGCAATTA | 1203 |
| DB | 3826 | AAAACAAAAAATTCACAACAATATTTAAAAACATTAATTTATTAATCATTTTCAAAAATATA | 3767 |
| QY | 1204 | GAAATATATTCGATGTGATGTGGAATATATGCGCTGAAAAAATTTAGTATATATGCTGT | 1265 |
| DB | 3766 | AAAAATATTTTAAAAAAAATTTAAAAATTTTTTTTATTTTTTATTTAAATTTATTTAA | 3707 |
| QY | 1264 | TGCTACTGTATTAATAAGGACATGAATGAAATTTAAACACTATATATATGCAAAAAATCA | 1322 |
| DB | 3706 | TCATTTTAAAAAATATATTTTATTTAAAAAAAATTTCAAAATTTATATACATATTTACAT | 3647 |
| QY | 1324 | TTCAATAGATTTTATATACATTTAATTTGGTCATATGTCTTGGAGATCTAATATTTACATG | 1385 |
| DB | 3646 | TACATTTAAATTTTATATTTAAAAATTTTTTTTAAATTTTAAATTTTTCATTTTATATAT | 3587 |
| QY | 1384 | TACAAACAACATTCACGCAATATGTGCTTTGGACTAGAAATATAGAAAAAGGTAGAAATAT | 1443 |
| DB | 3586 | TCATTTATATATTTTTCACACAATATTTTAAACATTAATTTTATTTAATAAAAAATAT | 3522 |
| QY | 1444 | AAATACATTAATAGATCAACAACCTTAAGCTTTGTGAAGAAACCACTACTGTAAAAACACT | 1503 |
| DB | 3526 | AAATTCATTAACAATTCATATATATCATATATTTTATTACAC-----AAAAAATA | 3475 |
| QY | 1504 | GATATCATTTAGCAAAAAAAGCTAATGTATGAGAACTACATTTGCATATCTATTTCAATTT | 1563 |
| DB | 3474 | ATTTTATTTATATATATATATATTAACAAATTTATATTTTCAACATTTTAAACAAAAATATTT | 3415 |
| QY | 1564 | ATTACATGAGATATATCACTAGATTAAGCCATATCAACATATTTATCTTAGTCTATCAT | 1623 |
| DB | 3414 | AAAAAAATTCAAATTTATTTATTTTTCATTTTATTTTATTTTAAATCATTCATTTCAAAATTTT | 3355 |
| QY | 1624 | CCTTTGATTAATTTCAACAATGTATATTAAGGTCAAAAAATATAGATTCACAAGTTCA | 1683 |
| DB | 3354 | ATTCATTTTTTTACATTTACATTTATATTTTTTTTTTTTTCABAAAAAATTAACAACAATA | 3295 |
| QY | 1684 | CCACATATATAAAAAGCAATCAATACCCATCGTTAACATTCACAATAAGTATGTACATA | 1743 |
| DB | 3294 | ACAAAAATTTTTTATACATCAATTAATTAATAAATAATTAATCAAAATTTTAAAAAATA | 3235 |
| QY | 1744 | TTTATATACAAACCTTAAGTATATCTTAAATTTCCGTGCAATATATATTAACAATTCCT | 1803 |
| DB | 3234 | TAAATAATTTAAATTTTATTTATTTTACATTTCAATTCACAACAATAATTTTATTTTAAAAATATATTT | 3176 |
| QY | 1804 | ATCTCCATCTCATTTACCAATTTGTATTCAGATTAATTAATTAAGTCAAAATTTAACTTTTA | 1863 |
| DB | 3175 | ATTTAAATTAATATATTAATAAACAATTTTTCATATCATTTAAATCATTTTAAAAAATTA | 3116 |
| QY | 1864 | CTTATTTTATTAATAATATGTGTGACACATTTTTCACATCATGTATATTAAGAATCAT | 1923 |
| DB | 3115 | ATTTTATATTAATAATTAACAATAAATAAACAATAATTTTCAATTTTCATTCACAATTTATTAAC | 3056 |
| QY | 1924 | AAATTTACATTAATAATTAATTAATAAAGAACCAATACATATATTAATTAATTAATTAACAC | 1983 |
| DB | 3055 | ATTTATTTTATTTTATTTATTTATTTATTAATAAATTTAAACATCATTTTAAAAATATATCT | 2996 |
| QY | 1984 | TATTTACTACGTATATATTAATAGATGTCTGCATATAATTTCTATCATCTCTGACAT | 2043 |
| DB | 2995 | AAATTTTCAATTAATTTTTTTTTTATTTTTTTTAAACAATATATATTCATTTAAATTTTACATTT | 2936 |

QY 2044 ATTCAACATATAAATATGATTAACGTGTTTTTATGCCCCATTAGTCTTAATAATTC 2103
DB 2935 TCATTAATAATTTTACATTTTTCATTTTCATTTTATTTTATTTTTCACCAAAATTT- 2877
QY 2104 CTAAATACCCAAAAGGTTAAATACACATCTCTGCACACATATGAGAGATACAAATAC 2163
DB 2876 -TTTTTCATTAATAATTTTCAAAATTTTATTAATTTTAAATAAAATATCATTTATTATTA 2818
QY 2164 AGCTCAATATAAATTAACAAATTTCTTAATATACCACTCTCTTAACATAGAACCAAT 2223
DB 2817 ATACACATTTTAAATAACAAATTCATTAATTTTAAATAAATAACATTAATTAATTT 2758
QY 2224 AGTATATCATACCTAAGATATATCCCTACAGCATATCAAAATATTAAGAAATATCG 2283
DB 2757 TTTTATTTTATTAATAATTTATTTATACAAATTAATTTT-----ATTAAAAAATAA 2703
QY 2284 TTATGATATTTTATACATTTACAAATTTCCAACTTCACAAACAAACCTCAAGTTCTAT 2343
DB 2702 TAATTAATTTTAAATAAATTTTAAATAATTTATTTTAA-----AATTTTAAAT 2654
QY 2344 TATTAACAAATTCAGAAATTAATATATCTTACCAACACATAGTTATCAAAATCTAATA 2403
DB 2653 TTTAACAATTTTATTTTATTTTATTTTATTAATTTATTAATTTTATTAATTTTATTTTCA 2594
QY 2404 CATCATCTCTTTTAAAGATGACTCAATTCATATCATTAATTAATTAATGCA 2463
DB 2593 TTATTTTATTTTATTTTAAATAATTTTATTAATTTTAAATTTTATTTTATTTATTA 2534
QY 2464 TGTTTATTCAGAAACATTCACATCCAGAAATTTCAATTTAGCGTATAGTGAAG 2523
DB 2533 ATTTATTTATTTTATTTTATTTTATTTTATTTATTTATTTATTT-----ATTAAAAAATAA 2478
QY 2524 TTCCATTCACAAATTCAGCGTTGATATCAATTAATCTCCAAAACGAGATACGAGT 2583
DB 2477 AATTAATAAATATTAATAATTTTAAATAAATAAATAATTTTAAATAAATAATTTATTT 2418
QY 2584 TTTTAAATATCTGCTTCAATCTCAGAGCTTACGTAATGAGACAAATTAATTT 2643
DB 2417 TTTTATTTTATTTTCAAAAAATTAATAATTTTATTAATTTTATTAATAAATAAATAATTT 2358
QY 2644 AGCATTAATAACATCTTAATAGCGAATTAATATCTATATCTAGTAAATCTTTTC 2703
DB 2357 AATTTTAAATAATTAATAATTTT-----TATTTAATAACAAATAATTAATAATTTTATA 2303
QY 2704 TGTAATCATCTCTGAGATTTCTATTAATAATTAATTAATTAATTAATTAATTAATTA 2763
DB 2302 AATAATTAATAATTAATAATTTTATTTTATTTTATTTTAAATAATTAATAAATAA 2243
QY 2764 TCTTAAGTACCTGGCATTCATATGATGATGCTATTAAGCTTCATCACTTAATAC 2823
DB 2242 TTTTATTTTAAATAATTTTATTTTATTTTATTTTATTTTATTTTATTTTAAATTTT 2183
QY 2824 ATAGATACCAACCATTTTATTTTCAAGATGATCAATAAATCCACACATTCGCG 2883
DB 2182 AATAATAATTAATAATTTTATTTTATTTTATTTTATTTTAAATAAATAAATAATTA 2123
QY 2884 TATATCTGATCAATTTCTATTAAGACGATCTCAATTTGAGAGACATTTATCTATA 2943
DB 2122 AATTTATTAATAATTTATTAATTTATTTATTAATAATTTTAAATAATTTTATTAATAAT 2063
QY 2944 ATATTAATCTCGCAAAACCTACGCTTTTACTTCAATTTCAATATCTTTGGAAGAGA 3003
DB 2062 TTTTAAAAAT-AAAAAATTTTAAATAATTTATTTTAAATAATTT----- 2013
QY 3004 CACAAATATCTCAGATCTATACATCAAGAAATCTTACTATTTTCCATATATTTTTC 3063
DB 2012 -AAAAATTTATTAATAATTTTAAATAAATAATTAATAATTTTATTTTAAATAATTTATA 1954
QY 3064 CTGCTGATTAATCTGAGTATCTTAATGCGATGATGTAATCTGTTACTTCATC 3123
DB 1953 TTTTATTTATTAATAATTAAT-TTAATTAATTTTATTTTAAATAATTTTATTTTAT 1895
QY 3124 TATTAATAACGATCATGATTTGCTGCTATTAATTCATTAATGAGAGTACTCTATG 3183

DB 1894 AATTAATAAATAAATAAATAATTAATTAATTAATTTTATTTTATTTTATTTTATTTT 1835
QY 3184 AGACAAATATTAATTAATTTCTCTCCTACCAAAATATTAATTAACCAAGATTAATTAAGCTG 3243
DB 1834 AACAATAATTTTATTTATTTATTAATAAATAAATAAATAAATAAATAAATAATTTTAAATTT 1775
QY 3244 TAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3303
DB 1774 T---TATTTAATAAATTTTATTTTATTTTATTTTATTTTATTTTAAATAATTAATAAATA 1719
QY 3304 TAGCTTACGACATCTCTCTTCCAAACCTGCGCAACGATATGATATACGACTCGA 3363
DB 1718 TATTAATTAATAATTTTAAATAATTTTATTAATTTTAAATAAATAAATAATTTATCATATTAAT 1659
QY 3364 TGATAGCTGCTTATTTATTCATCTCTCAATTAACGATTAACAGGGAGTTACCGAC 3423
DB 1658 TAATAATTTTATTTTATTTATTTTCAATTTCAATTTTATTTTAAATAATTAATAATTT 1599
QY 3424 TATTTCAATAATCTCAAAATGCTGATTTCAATCATATTTTCTTACGATATCTACAG 3483
DB 1598 ATAAATTAATAATTAATAATTTTAAATAATTAATAATTTTAAATAAATAATTTTATTTA 1539
QY 3484 CTCTATGCTCTCTTTACAACTTAATCTTCTCTCTGTAAGGCTTTCTATTA 3543
DB 1538 AAAAAATTTTAAACATATTTTAAATAATTTTATTAATTTTAAATAATTTTAAATAATTTA 1479
QY 3544 ATCATTAAGCACCAATATGACAGACTTTACAGAGTGGCAATTAACAATGCCACTAAT 3603
DB 1478 TTTTAAATATATTTTATTAATAAATAATTAATTTTATTTTATTTTAAATAAATAAATAA 1419
QY 3604 CATTAATACGGAATAAGATATCTTCTTTAG-----CTTTCCGATGACG 3654
DB 1418 AAAAAATTAATAAATAAATAAATAATTTATTTTAAATAAATAAATAAATAAATAAATAA 1359
QY 3655 TAAATCATATATATGATCTCTTAACATATATCCAAATATTAACAAATACAGGCTCTT 3714
DB 1358 TTAATAATTAATTTTATTTTATTAATAAATAATTAATAAATAATTTTATTAATTAATTT 1299
QY 3715 TTTCAAGCATCTTGATCGCGGATTAACCATGATCACTGCTAATTTAGTACATTAATTC 3774
DB 1298 TTTTAAATAATTAATAATTTTATTTATTAATTAATAATTAATAATTTTATTTATCA 1239
QY 3775 ATCACTTAATATATCTTTATTTAGATTT-CTGATATCACTTATCATATCAACATTAATA 3833
DB 1238 AAAAAATTTTATTTATTTAATAAATAAATAAATAAATAAATAAATAAATAAATAATTT 1179
QY 3834 CTTCAGAAATTAACAATCTTCTGGAATTTCAAAATCTGCTGCATTAATTAATCTACA 3893
DB 1178 AATTAATATCAAAATAATTAATAATTTTAAATTTTCAAAATATCAAAAAATAAATAAATA 1119
QY 3894 ATACTTAATGAATAATGCAACACAAATGATCTCTTAAAGTACCTCAATTAATTAATCTAC 3953
DB 1118 CAATATATTAATTTTAAACATTTAAATAAATAAATAAATAAATAAATAAATAAATAATTT 1059
QY 3954 ACACACATATGATCAACAACAATTAATTAATTAATTAATTAATTTTATCTTAATAATTT- 4012
DB 1058 TTAATATTAATAATTTTCAAAATTAATAAATAAATAAATAAATAAATAAATAAATAATTTT 999
QY 4013 -----AAACATTAATAAATAATTAATAATTAATTAATTAATTAATTAATTAATTAATTA 4067
DB 998 AAAAAATTTATTAATTAATTTTATTTATTTATTTATTTAATAAATAAATAAATAATTTTAA 939
QY 4068 CTATATATCCAAAAGAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAATTT 4127
DB 938 AAAAAATTTATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 879
QY 4128 ACCATTA-GTATCAATTAATTAATAATTTTCAATTAATTTTCAATTAATTTTCCCTGGAACAG 4186
DB 878 TTTTATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT 819
QY 4187 CATGCTATAGTTTCCATTAATAAATACTGAAAGCTTTATTAATTAATTAATTAATTC 4246

Db 547 GTGGGTGTAACCTAGAGAGCATTAAAAATATATATTCAGCTGCAAGTGTATTGTA 606
 Qy 1235 GCTGGAATAAATTAGTAATATATGCTGTGCTACTGTATATTAAGCAATGATGAA 1294
 Db 607 GCAGGCGATGGGCTATGCTATATGCTAAGGCTTTATGACAGAGTTAGCTGAA 666
 Qy 1295 ATTAAACCTATATATAGCAAAAAATCATTCATATGATCTTCATATTAATTTGTCCA 1354
 Db 667 ATAGCCGATATAGTAAAAAATTAGTGTCT-----GATCAATGACATTTTATGCTCT 720
 Qy 1355 TCATGTCTTGAGATCTAATATATTAATCATGATGACAGCAATTCAGCAATATGCTTTT 1414
 Db 721 GGTGTATTGAGATCTAATTTGTTACTGTACATCAAGCATTCACGTATATACCTTTA 780
 Qy 1415 GGACTGAATATAGAAAGAGTAAAAATATTAATACCTTATATGATCAACACTAAAGCTT 1474
 Db 781 GGAATTAATATAGCAAGGCAAGCAAGATGAAAGCTTTAAATGAATGAATATGGA 840
 Qy 1475 GTTGAAGAACCACTACTGTAAAAACCACTGATATCATTTAGCAAAAAAATTATGTAGAA 1534
 Db 841 GTAGAGGAGATTTATATCTATCTATCTGTATACCATTTAGCTAAACAAATATGAT 900
 Qy 1535 CTACCAATTGCAATATCTATTTTCAATTTATTTATCATGAGATA 1577
 Db 901 ATGCCCTATATCTAATGCACTTTACAGGTTTATTCGAGATA 943

RESULT 10

ACF75135
ID ACF75135 standard; DNA; 996 BP.

ACF75135;

DT 20-NOV-2003 (first entry)

XX Staphylococcus aureus DNA #2815.

KW Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;

XX enzymatic assay; antibiotic target; gene; ds.

OS Staphylococcus aureus.

XX WO200294868-A2.

XX 28-NOV-2002.

XX 27-MAR-2002; 2002WO-IB002637.

XX 27-MAR-2001; 2001GB-00007661.

XX (CHIR-) CHIRON SPA.

XX Masignani V, Mora M, Scarselli M;

XX WPI; 2003-120786/11.

XX P-PSDB; ABM73575.

XX New Staphylococcus aureus protein, useful as a vaccine for treating or

XX preventing Staphylococcal infection, specifically an infection caused by

XX S. aureus, e.g. sepsis.

XX Claim 6; SEQ ID NO 5629; 49pp; English.

CC The invention relates to novel genes and encoded proteins from
 CC Staphylococcus aureus. A composition comprising the S. aureus protein, a
 CC nucleic acid encoding the protein, or an antibody to the protein, is
 CC useful as a pharmaceutical, particularly as a vaccine for treating or
 CC preventing infection due to Staphylococcus bacteria, specifically an
 CC infection caused by S. aureus. The composition is particularly useful for
 CC treating or preventing sepsis in a patient. The composition can also be
 CC used for diagnostics. The protein is also used in an assay for enzymatic
 CC studies and as a target for antibiotics. This sequence represents one of
 CC the novel S. aureus genes of the invention

XX Sequence 996 BP; 345 A; 134 C; 210 G; 307 T; 0 U; 0 Other;
 SQ Query Match 3.5%; Score 168.6; DB 7; Length 996;
 Best Local Similarity 51.4%; Pred. No. 1e-17;
 Matches 505; Conservative 0; Mismatches 454; Indels 24; Gaps 4;

Qy 635 AAAATCGATTTTATAGTGAAGATCATTTGGCAGCAATATGCAATGTCATGTCAGCA 694
 Db 7 AAAATTCGCTTTTGTATAGGAGGTTTGGACAGCCCTTCCCAATGTTCTTGGAGAA 66
 Qy 695 CANGATATCAGTAACTTATGAGGAGCGATCATAGAAATATACATATATAACACT 754
 Db 67 AATGACATGATGTTTATGATGTTGGGTAATAATCAAGATGCTGTATGATTAATTAACA 126
 Qy 755 TACGAAAAAATTTAAATTTTAAACCATATCATCTACAGACAAATATATGCAAC 814
 Db 127 TGTCTATCAAAATATAAAGTATTTAAATACGCAAAATTAGATGTTAATCATATGCTACT 186
 Qy 815 AGCAATATAGACGAGATATATCTGACAAATATATGATATATTTACTATCTCTACA 874
 Db 187 TCGATATGACCAAGGATATCAATTTGCGATATTTACTTAATG---CTTACCTACT 243
 Qy 875 CAACAATTAGCACCATATGTATACAAATATCAACAAACAGCATATGTGTAAAAATACT 934
 Db 244 AAAGCAATGGAGAAAGTGTCTCTCAAAATATATATATAGC-----TGACCTTAATAAG 297
 Qy 935 CCAATATTAATTTGTATAGGATATGCAATATATCACTCAATATTTCCAGTGAATA 994
 Db 298 ACTTTTATACATGTTGCTAAGGATATGGAATGGAACGTTTAAAGTGTATAGAAATG 357
 Qy 995 GCAGAGAAAAATTTT-----ACAATATATCCAAATTTTATCTCTGTGTCAGT 1045
 Db 358 ATGGAATATCTATATTCACCTGAATATATAGCAGTATTTGGCTGTGTAGGCGCAAGT 417
 Qy 1046 TTGCTTAAGAAATATGAGAAACATCTTCCCTGTATATATGATATGCTGTGATATTA 1105
 Db 418 CATGCGGAAGATTTGATGCAAGCAACCACTACAGCTTCTCTCATATAAAGATTA 477
 Qy 1106 GAACCTGTGATATCATATGATGAGAAACATATGATATGATGTTCTAAAAATATATCAAT 1165
 Db 478 AGTATAGTAATATTAAGCAAGATTTATATGATATATTTGCGTGTATACAGAAAT 537
 Qy 1166 CAAGATATTTATAGTGTACAGATTTGAGCTGCTTAAGAACAATATGCAATTTGCAATG 1225
 Db 538 GATGACTGTATGTTGTTGAACCTTGTGTGACATGAAATATATCATCGACATGCAAGT 597
 Qy 1226 GGAATATCGCTGAAAAAATTTAGTATATATGCTGTTGCTACTGTTATATATTAAGGC 1285
 Db 598 GGTATCTAGCTGGAATTTGGCTACGGGTATATGCAAAACCTCATTAATGACTCGTGGC 657
 Qy 1286 ATGATGAATAATTAATACTATATATAGCAAAAAATCATTCATATGATGCTCATATAC 1345
 Db 658 TTACGGAATAATTTAGTA-----TTAGTGAATAAGTTAGTGCCATCTATAGCAATTT 711
 Qy 1346 ATTGTCATCATGCTCTTGAGATCTAATATTAATATGTAACAGTAACAGAACTTACGCAAT 1405
 Db 712 CTAGCTTTAGTGTATGCGTATGACTTATATGCTTATCTGACATCAACACATTCGGAAT 771
 Qy 1406 ATGGCTTTTGGACATGAAATAGGAAAAAGTGAATAATATATATCATTAATATATGCAAC 1465
 Db 772 TTACATATAGATATTAACCTTGACAAAGTGAATCAATGATCAAGCATTAATCTGAATG 831
 Qy 1466 CTAAAGCTTGTGAGAGAACAGTATGTAATAACCTGATATATCATTTAGCAAAAAAATCT 1525
 Db 832 AATATGTTGTGTAAGTATTTATACAACTAAATCATGATTTATCATTTAGTAAAGAAAA 891
 Qy 1526 AATATGAACATCAATTTGCAATATCTATTTAATTAATTAATGAGAAATATATACATA 1585
 Db 892 AATGTGATATGCAATTAACAAATGCAATTAATATATGATATTTATGAAATATCTCAGTA 951
 Qy 1586 GATTAAGCCATATCAACATATTT 1608

Db 952 AAGAAATGCGTAAAGATTAAAT 974

RESULT 11

ID AAS55337 standard; DNA; 999 BP.

XX AAS55337,

DT 13-FEB-2002 (first entry)

DE Staphylococcus aureus DNA for cellular proliferation protein #1649.

XX Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;

XX Staphylococcus aureus.

XX MO200170955-A2.

XX 27-SEP-2001.

PF 21-MAR-2001; 2001WO-US009180.

XX 21-MAR-2000; 2000US-0191078P.

PR 23-MAY-2000; 2000US-0206848P.

PR 26-MAY-2000; 2000US-0207727P.

PR 23-OCT-2000; 2000US-0242578P.

PR 27-NOV-2000; 2000US-0253625P.

PR 22-DEC-2000; 2000US-0257931P.

PR 16-FEB-2001; 2001US-0269308P.

XX (ELIT-) ELITRA PHARM INC.

PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GT,

PI Yamamoto RT, Xu HH,

DR WPI; 2001-611495/70.

XX P-PSDB; AAU37478.

XX New polynucleotides for the identification and development of

XX antibiotics; comprise sequences of antisense nucleic acids.

XX Claim 27; SEQ ID NO 8974; slipp; English.

XX The invention relates to antisense inhibitors of genes essential to

XX prokaryotic cellular proliferation, their use in identifying the genes,

XX themselves and the encoded proteins. The prokaryotes used are Escherichia

XX coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,

XX Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also

XX useful for the identification of potential new targets for antibiotic

XX development. The antisense nucleic acids can also be used to identify

XX proteins used in proliferation, to express these proteins, and to obtain

XX antibodies capable of binding to the expressed proteins. The proteins can

XX be used to screen compounds in rational drug discovery programmes. The

XX antisense nucleic acid sequence is also useful to screen for homologous

XX nucleic acids which are required for cell proliferation in a wide variety

XX of organisms. The present sequence encodes an essential prokaryotic

XX cellular proliferation protein. Note: The sequence data for this patent

XX did not form part of the printed specification, but was obtained in

XX electronic format directly from WIPO at

XX ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 999 BP; 347 A; 134 C; 210 G; 308 T; 0 U; 0 Other;

Query Match 3.5%; Score 168.6; DB 4; Length 999;

Best Local Similarity 51.4%; Pred. No. 1e-17;

Matches 505; Conservative 0; Mismatches 454; Indels 24; Gaps 4;

QY 635 AAAAAATGATTTTGGGAGGATCATTTGGGAGGAAATAGCAATATGCTGTACGA 694

XX AAAAAATGATTTTGGGAGGATCATTTGGGAGGAAATAGCAATATGCTGTACGA 66

QY 695 CATGATATATCAGTTAACTTATGAGGACGATCATAGAAATATTAACAATATAACACT 754

Db 67 AATGACATCATGTTTGGATGAGGAAAAATCAAGATGCTTGTATGAATTAATAACA 126

QY 755 TACGAAAAAATTTAAATATTTTACCACATATCATCTACAGACAACATATATGCAACC 814

Db 127 TGTCAATCAATATAAAGATTTTAAATAGCGAAATATGATGTAAACATCATCGCTACT 186

QY 815 AGCAATATAGACGAGATATATCTGACAAACATATATATTTCTTAATCTTCTCTACA 874

Db 187 TCGATATGACCAAGAAATTCATTTGCGATATTTACTTAATGG---CTTACTACT 243

QY 875 CAACAATTAAGCAACATATGATACAAATACAAACAAACAGCATATGTATTAATAACT 934

Db 244 AAGCAATGCGAAGAGTGTCTCTCAATTAATATGATAGC-----TGACCTTAATAAG 297

QY 935 CCAATATTAATTTGTATGTAAGATGATGCAATATATATCTCAATTTCCAGTGAATA 994

Db 298 ACTTTATACATGTTGCTTAAGGATTTGAATAATGGAACGTTTAAAGCTGTGCAAAAG 357

QY 995 GCAAGAAATTTT-----ACATATTAATCAATTTTATCTCTGTGCAAGT 1045

Db 358 ATGGAAGATTCATTTCACTGAAATATATGACAGTATTTGGGCTGTGTACAGGCCAAGT 417

QY 1046 TTGCTTAAAGAAATTTGCAACATCTTCTCTGTATATGATCTGTGCTGTATATTA 1105

Db 418 CATGCGAAGAAATTTGATGCAAGCAACCACTACAGTGTGCTTCACTCAAAAGATTA 477

QY 1106 GAACCTTGATCATTTGATATAGAAACATATATGATATGTTTAAATAATATATACAT 1165

Db 478 AGGTAGATTAATTAAGCAAGATTTATTAATGATATATTTGGTGTATGACAGAT 537

QY 1166 CAGATATTTATGATGATACATTTGAGAGCTGCTTAAAGACATATTTGCAATGATGT 1225

Db 538 GATGACTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 597

QY 1226 GGAATATCGCTGGAAGAAATTTAGATTAATATGCTGTGCTACTGTATTAAGGC 1285

Db 598 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 657

QY 1286 ATGATGAATTAATAACATATATATATATATATATATATATATATATATATAT 1345

Db 658 TTAGCGAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 711

QY 1346 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1405

Db 712 CTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 771

QY 1406 ATGCTTTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1465

Db 772 TTACATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 831

QY 1466 CTAAGTGTGTAAGAAACAGTACTGTAATAACCTGATATCAATTAGCAAAAAAATT 1525

Db 832 AATATGTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 891

QY 1526 AATGTAAGTATGCAATTTGATGATGATGATGATGATGATGATGATGATGATG 1585

Db 892 AATGTAAGTATGCAATTTGATGATGATGATGATGATGATGATGATGATGATG 951

QY 1586 GATTAAGCAATATCAACAATATT 1608

Db 952 AAGAAATGCGTAAAGATTAAAT 974

RESULT 12

ID ACA27525 standard; DNA; 999 BP.

XX ACA27525;

DT 19-JUN-2003 (first entry)

WP AAX20248_05 500001 610000
 WP AAX20248_06 600001 710000
 WP AAX20248_07 700001 810000
 WP AAX20248_08 800001 910000
 WP AAX20248_09 900001 910715
 ID AAX20248 standard; DNA; 910715 BP.
 AC AAX20248;
 XX
 DT 04-MAY-1999 (first entry)
 DE Borrelia burgdorferi polynucleotide sequence #1.
 XX
 KM Borrelia burgdorferi; spirochete; bacterium; pathogen; Lyme disease;
 KM epidemic relapsing fever; endemic relapsing fever; Lyme borreliosis;
 KM infection; diagnosis; characterisation; detection; ds.
 OS Borrelia burgdorferi.
 XX
 EN MO9858943-A1.
 PD 30-DEC-1998.
 XX
 PF 18-JUN-1998; 98MO-US012764.
 XX
 PR 20-JUN-1997; 97US-0050359P.
 PR 22-JUN-1997; 97US-0053344P.
 PR 22-JUN-1997; 97US-0053377P.
 PR 03-SEP-1997; 97US-0057483P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PA (MEDI-) MEDIMUNE INC.
 XX
 PI Fraser C, White OR, Clayton R, Dougherty BA, Lathigra R;
 PI Smith HO;
 DR WPI; 1999-081217/07.
 XX
 PT New isolated Borrelia burgdorferi nucleic acids - used to develop
 PT products for the detection, diagnosis, characterisation, prevention and
 PT therapy of infections, particularly Lyme disease.
 XX
 PS Claim 1; Page 157-671; 1128PD; English.
 CC AAX20248 to AAX20402 represent polynucleotide sequences isolated from
 CC Borrelia burgdorferi (Bb). Products derived from Bb can be used for the
 CC detection, diagnosis, characterisation, prevention and therapy of Bb
 CC infections, e.g. Lyme disease. They can also be used for the production
 CC of biosynthetic products, e.g. enzymes. Borrelia belongs to a family of
 CC motile, spiral-shaped bacteria called Spirochetes. Spirochetes are
 CC pathogenic in humans and Borrelia causes epidemic and endemic relapsing
 CC fever, and Lyme borreliosis, more commonly known as Lyme disease
 CC
 SO Sequence 910715 BP; 327171A; 129646C; 130753G; 323091T; 0U; 540ther;
 Query Match 3.5%; Score 165.8; DB 2; Length 110000;
 Best Local Similarity 49.4%; Pred. No. 3.4e-17;
 Matches 598; Conservative 0; Mismatches 577; Indels 36; Gaps 5;
 QY 2661 TAAATAGCGAATTAGATCTATATCTACTGTAATCTTTGCTGTAATCATCTCCTTAG 2720
 DB 106138 TAAGATTTTCAATTATTTTATTTTGTCTGCAAACTT---CTTAGGATTTTGTTCG 106082
 QY 2721 GAGATTTCAATTAATCAATTCATTAATCAATTAATCAATTAATCAATTAATCAATTA 2780
 DB 106081 CTTTGATTAATTAATTAATCAATTCATTAATCAATTCATTAATTAATTAATTAAT 106022
 QY 2781 ATTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2840
 DB 106021 AATGATTAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 105962
 QY 2841 TTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2900

DB 105961 TTTCTTTGCGACGCTTTTATTAATGATGATGATGATGATGATGATGATGATGATG 105902
 QY CTCTAAGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2901
 DB 105901 CTCTAAGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 105842
 QY AACTAAGAGGCTTTTACTTCACTTTCATTAATGATGATGATGATGATGATGATGATGATG 2961
 DB 105841 ATTTTCTTTTATTTATTTCCCTTCAATGATGATGATGATGATGATGATGATGATG 105782
 QY CTATTAAGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3021
 DB 105781 CACTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 105722
 QY 3081 GTAATCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3140
 DB 105721 AAGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 105662
 QY 3141 GATTTGCTGCTTAATTAATTCATTAATGATGATGATGATGATGATGATGATGATGATG 3200
 DB 105661 TGTTCGAAGTCAATCTGCTTTTGTGGAATGCTCCAGTGAATGCT-----C 105608
 QY 3201 TTTCTCACTACCAATATATTAATTAATGATGATGATGATGATGATGATGATGATGATG 3260
 DB 105607 CTTTTCATGCGCAAAAGTCACTTCAAGAAATGCTTCAAGAAAGTGGCGCAATTTA 105548
 QY 3261 TAGATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3320
 DB 105547 CTTTATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 105488
 QY 3321 CCTTCCACACCTGGGGAACAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 3380
 DB 105487 CTTTACCAACACCGCTTTCGCGCTTAATGATGATGATGATGATGATGATGATGATG 105428
 QY 3381 TTAATCAATCTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3440
 DB 105427 TTAATCAATCTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 105368
 QY 3441 AT-GGTGATTCATCATCATCTTCTCTAGTATCTACCAAG----- 3483
 DB 105367 TTAATCTTCTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 105308
 QY 3484 ---CTTAATGCTCTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3539
 DB 105307 TTTTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 105248
 QY 3540 TATATCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3599
 DB 105247 AAAAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 105188
 QY 3600 TAAATCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3659
 DB 105187 TTAATCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 105128
 QY 3660 CATCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3719
 DB 105127 CAGATATGCGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 105074
 QY 3720 AAGCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3779
 DB 105073 CAATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 105014
 QY 3780 TTAATTAATCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3839
 DB 105013 CAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 104954
 QY 3840 AATTAATCAAT 3850
 DB 104953 TCAATTTCAAT 104943
 RESULT 14
 AAX20248_01/c

CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-regulated gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC prokaryotic essential genes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WFO at
 CC ftp.wipo.int/pub/published_pcr_sequences
 CC XX

Sequence 999 BP; 349 A; 133 C; 210 G; 307 T; 0 U; 0 Other;

Query Match 3.5%; Score 165.4; DB 7; Length 999;
 Best Local Similarity 51.2%; Pred. No. 3.2e-17;
 Matches 503; Conservative 0; Mismatches 456; Indels 24; Gaps 4;

QY 635 AAATACAGTATTTAGTGTGAGATCATTTGACAGACAGATGCAATTCGCTGACGA 694
 DB 7 AAAATACCGTTTGGTATGAGAGTTTGGACAGCCCTGCGCAATGTTCTTGAGAA 66
 QY 635 CATGTATATCACTTATGAGAGTGTATGAGATTTATTAACATATTAACACT 754
 DB 67 AATGACATGATGTTTATGATGGGGTAAATCAAGTGTGTGATTAATTAACA 126
 QY 755 TACGAAAAAATTTAAATATTTTACCACATATCATCTACAGACACATATATGCAAC 814
 DB 127 TGTCTACAAATTAAGATTTTAAATACCGAAATTAAGTTTACATCATGCTACT 186
 QY 815 AGCAATATGACGAAGTATATCTGACAAACATACATGATATTTACTATCTTCTACA 874
 DB 187 TCAGATATGACCAAGCAATTCATTTGAGATATTTTCTTAAGG---CTTTACTACT 243
 QY 875 CAACATATACGACCATATGTAACAATAACAAACAGCATATGTGTAATAATCT 934
 DB 244 AAGCAATGCGAAGATGCTACTCAATTAATGATTAAGC-----TGACCTTAATAAG 297
 QY 935 CCAATATTAATTTAGTAAAGATTCGAAATTCATCATCAATTTCCAGTGAATA 994
 DB 298 ACTTTATACATGTTGCTAAAGGTATTAAGAAATGAGAGCTTAACCGTGTGCAAGATG 357
 QY 995 GCAGAGAAATTT-----ACAATATATCAATTTTAACTCTGCGGCCAAGT 1045
 DB 358 ATTGAAGATCTATTTACCCGAATATATGACAGTATGCGCGTGTCTCAGGCCAAGT 417
 QY 1046 TTGCTAAAGAAATTCGAAACATCTCTTGTATATAGTACTGCTGTGATTAATAA 1105
 DB 418 CATGGGAGAAAGTGTAGTCAAGCAACCACTACAGTGTGATCATCAAAAGATTA 477
 QY 1106 GAATTTGGATCTATTGATGAACCAATTAAGTATGATGTTCTTAATAATATATACAT 1165
 DB 478 AGTGAAGTAAATTAACGCAAGATTTATTTATGATGATTTATTTGCGGTGTACACAA 537
 QY 1166 CAAGATATTAAGTGTACAGATTGAGCTGCAATTAAGAAACATTAATTCATTTGATGT 1225

DB 538 GATGATTTATGTTGTTGAACCTTGTTGATGATGAAATAATCATGCACTAGCAAGT 597
 QY 1226 GGAATATATGCTGGAAAAAATTTAGTATATATGCTGTTGCTACTGTTATTAACAAAGC 1285
 DB 598 GGTATGTAGCTGGAATTTGGCTACCGTATTAATGCAAAAGCTGCTTAATGACTCGTGT 657
 QY 1286 ATCAATGAATTAATAACATATATATAGCAAAAAATCATCAATGATCTTCAATCACTA 1345
 DB 658 TTACCGAAATTAATGAGA-----TTAGTGAAAGATTAAGTCCGATCTTATGACATTT 711
 QY 1346 ATTTGTCATCATGCTTGTGAGATCTAATATTAACATGTAACAGCAACATTCACGCAAT 1405
 DB 712 CTAGGTTTATGTTGATGATCGTACTTAATCGTTACTTGTGATATCAACATTCCTGAAAT 771
 QY 1406 ATGGCTTTTGAAGTAAATAGAAAGTAAAGTAAATATTAATATATATATATAGTACAAAC 1465
 DB 772 TTACATTTAGATATTAATTAATTTGACAGAGTGAATCAATGATCAAGCATTTATCTGAATG 831
 QY 1466 CTAAAGCTTTGTAAGAAACAGTACTGTAACCACTGATATCATTTAGCAAAAAAATCT 1525
 DB 832 AATATGTTGTAAGATTTATTAATCAATTAATCAATTAATCAATTTAGCTAAAGAAAA 891
 QY 1526 AATGTGAACCTACCAATTTGATATCTAATTTATTAATTAATTAATTAATTAATTAATTA 1585
 DB 892 AATGTGAATATGCAATTTCAAAATGATATATTAAGATTTATTTGAAATATTCACGTA 951
 QY 1586 GATTAAGCCATATCAAAACATATT 1608
 DB 952 AAAGATGCGTAAAGATTTAAT 974

RESULT 16
 ABZ10246
 ID ABZ10246 standard; DNA; 8056 BP.
 XX
 AC ABZ10246;
 XX
 DT 16-JAN-2003 (first entry)
 XX
 DE Haematopoietic cell proliferation disorder related DNA sequence #386.
 XX
 KW Human; haematopoietic cell proliferation disorder; cytostatic;
 KW Gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
 KW Cytosine methylation state; gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200277272-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 26-MAR-2002; 2002WO-BP003401.
 XX
 PR 26-MAR-2001; 2001US-0278333P.
 XX
 PA (EPIC-) EPIGENOMICS AG.
 XX
 PI Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;
 PI Olex A, Piepenbrock C, Adorian P, Grabs G, Lesche R, Liu B;
 PI Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T, Pelet C;
 PI Schwope I, Ziebarth H;
 XX
 DR MPI, 2003-018942/01.
 XX
 PT Detecting and differentiating between hematopoietic cell proliferative
 PT disorders, comprises contacting a target nucleic acid with a reagent that
 PT distinguishes between methylated and non-methylated CpG dinucleotides.
 XX
 PS Claim 28; SEQ ID NO 386; 117bp; English.
 XX
 CC The present invention describes a method for detecting and
 CC differentiating between haematopoietic cell proliferative disorders

QY 1957 ATACATATATTAATTAACAAACGATTTTACTAGCAATATATACATAGATGCTCGA 2016
 DB 2013 AATATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2072
 QY 2017 CATTAATTAATTCATCATCTTGACATATTTCAACATTAATG-----ATAAATG 2070
 DB 2073 AATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2132
 QY 2074 TTTTATTAATGCGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2130
 DB 2133 TTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2192
 QY 2131 CATTCCTGCACAAATATGAGAGATAGATACATACAGTCAATTAATTAACAAATTTCTA 2190
 DB 2193 AAAAAATGAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTT 2252
 QY 2191 ATATAACCCCTACTCTAAACATAGGAACCAATAGTATATCATCTAAGTAAATATCC 2250
 DB 2253 AATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2312
 QY 2251 CTACAGCATGTATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2310
 DB 2313 TATTAATTTTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTT 2372
 QY 2311 TTCCAACTTCAACAAACCTCAAGTCTTATTAACAAATTCAGAAATTAATTAATAC 2370
 DB 2373 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTAT 2432
 QY 2371 TTCACCAACATAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2427
 DB 2433 TTTTAAATTAATTTTATTTTATTTTATTAATTAATTTTATTTTATTTTATTTTAA 2492
 QY 2428 ACTTCACATTCATATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2487
 DB 2493 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2552
 QY 2488 TCACAAATTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2547
 DB 2553 TTTTAAATTAATTAATTAATTTTAAATTAATTAATTAATTAATTAATTAATTA 2612
 QY 2548 TCGTATATCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2607
 DB 2613 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2672
 QY 2608 TTCAGAGCTTACGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2667
 DB 2673 ATTA--ATTTTAAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTT 2730
 QY 2668 CGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2714
 DB 2731 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTATTT 2777

RESULT 17
 ADA31919/c
 ID ADA31919 standard; DNA; 1557 BP.

AC ADA31919;

DT 20-NOV-2003 (first entry)

DE DNA encoding *Acinetobacter baumannii* protein #3206.

XX de; gene; *Acinetobacter baumannii*; bacterial disease; antibacterial;
 KW vaccine; plant biocontrol agent.

OS *Acinetobacter baumannii*.

XX US6562958-B1.

XX 13-MAY-2003.

XX PD

PF 04-JUN-1999; 99US-00328352.
 XX 09-JUN-1998; 98US-0088701P.
 PR (GENO-) GENOME THERAPEUTICS CORP.
 PA Breton G, Bush D;
 PI WPI: 2003-576092/54.
 DR P-PSDB; ADA36045.
 PT New *Acinetobacter baumannii* proteins and nucleic acids, useful as reagents
 PT for diagnosing a bacterial disease, as components of antibacterial
 PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
 PT plants.
 PS Example; SEQ ID NO 3206; 328bp; English.
 XX The invention relates to isolated *Acinetobacter baumannii* nucleic acids.
 CC The A. *baumannii* nucleic acids and polypeptides are useful as reagents
 CC for diagnosing a bacterial disease, as components of antibacterial
 CC vaccines, as targets for antibacterial drugs, to detect the presence of
 CC A. *baumannii* and other *Acinetobacter* species in a sample, in screening
 CC compounds for the ability to interfere with the A. *baumannii* life cycle
 CC or to inhibit A. *baumannii* infection, and as biocontrol agents for
 CC plants. The present sequence represents DNA encoding an A. *baumannii*
 CC protein.
 SQ Sequence 1557 BP; 481 A; 321 C; 330 G; 425 T; 0 U; 0 Other;
 XX
 Query Match 3.4%; Score 160; DB 8; Length 1557;
 Best. Local Similarity 48.3%; Pred. No. 2,4e-16;
 Matches 553; Conservative 0; Mismatches 575; Indels 18; Gaps 3;

QY 2699 TTGCTGTAATCATCTCTTACGAGATTTCAATTAATTAATTAATTAATTAATTAATTA 2758
 DB 1208 TCTTGTGATTAACCTGACGACGACGATTAATTAATTAATTAATTAATTAATTAATTA 1149
 QY 2759 AATGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2818
 DB 1148 AGCTGACGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1089
 QY 2819 AATACATGATCACAAACCTATTTTATTAATTAATTAATTAATTAATTAATTAATTA 2878
 DB 1088 AATATTTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1029
 QY 2879 TCGGTATATCTGTACAAATTTCTAATTAATTAATTAATTAATTAATTAATTAATTA 2938
 DB 1028 ATCGGATGCTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 969
 QY 2939 CTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2998
 DB 968 CGGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 909
 QY 2999 GAAGACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3058
 DB 908 GTTGCGGACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 849
 QY 3059 TTTCCCTCTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3118
 DB 848 TACCATCTGCGACGACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 789
 QY 3119 TCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3178
 DB 788 TCATCAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 729
 QY 3179 CTATGACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3238
 DB 728 GCAACCGTAA-----AGACCTTTTCAATTAATTAATTAATTAATTAATTAATTA 675
 QY 3239 GCTGTAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3298
 DB 674 TTAGGAATGCTGCATGTTTAAAGCAATTAATTAATTAATTAATTAATTAATTAATTT 615

| | | | |
|----|-------|-----------------------------------------------------------------|--------|
| OY | 924 | GTAAAAATPACCCAAATTTTATTTGTGTGAAGGATACGAAATTAATCATCACTCAAAATTC | 963 |
| Db | 44213 | AAAAAAATCTACCTCTCAGAAAAATCAATACCACTCATTAABAAAAATATAACATTAACCTAA | 44154 |
| OY | 984 | CCAAGTAATTCGAGAAAGAAATTTTACATATATAATCCAAATTTTATATCTCTCGTCCAA | 1043 |
| Db | 44153 | ATATCTCCCCATPACAAATATATTCACACTACATACAAACCTTTTAAACAAACAAAATATA | 440594 |
| OY | 1044 | GTTTGTCTAAAGAAATTCAGAAACATCTCTCTGT--AGATATGTACTTGCTGTGATATA | 1101 |
| Db | 44093 | ACTTAAAAATATAAATATTCAAACACACCTTATTCATATAAAAAATATAACCAACAAAAA | 440344 |
| OY | 1102 | TAAAGAACTTGTCATTCATTGATAGAACATATPAGTAATGATGTTCTTAAAAATATATA | 1161 |
| Db | 44033 | AAAAAATAAATTTATACCAACAAATATATATAATACCAAAAAACAAAATTTAAAAACATA | 439744 |
| OY | 1162 | CCATCAAGATATATATAGGTGTCAGATTTGAGCTGCATTTAAAGACATATATTCATTCG | 1221 |
| Db | 43973 | CATTAATAATTTTAAAAATTTACG-AAAAATCTCTCTTAAACAAACCTTTACACATTA | 439151 |
| OY | 1222 | ATGTGGAATATCGCTGGAATAAATTTTAGTAATATCTGTTGCTATATATACTAA | 1281 |
| Db | 43914 | AAAAATAATATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATCGTATAT | 438558 |
| OY | 1282 | AGCATGATGTAATTAATAACACATATATATAGCAAAATAATCATTCATATGATCTTCATAC | 1341 |
| Db | 43854 | AAAAAATAAATAACATTAATCTCAAAACAAAAAATAAATAAATAAATAAATAAATAAAT | 437958 |
| OY | 1342 | ATTAATGTGTCATCATGCTCTTGAGAGTCTATATTTACATGTACACAGAACATTCAAG | 1401 |
| Db | 43794 | AACAAATAATCTTTATTTATTTATTTATTTATTAACAATTAATCCAAATAAATAATTAAAA | 437358 |
| OY | 1402 | CAATATGCTTTTGAGCTAGAAATATAGAAAGATAGAAATATATAATATATAGATCA | 1461 |
| Db | 43734 | ATATATATATTTTAACTTTAA---ACTATTCATCATATTAACATTAACATTAATAACAATT | 436779 |
| OY | 1462 | CACCTTAAAGCTTTGTAGAGAACCAAGTACTGTAAACCATGATATCATTTGCAAAAA | 1521 |
| Db | 43678 | CTAATCTAAAAATCAAAAAATTTAAAAACGTAAATTAATAAATAAATACTTCTAAAAACAAAA | 436198 |
| OY | 1522 | ACTTATATGTAGATACCAATTTGSCATATCTATTTCAATTTATATATCATGAGAAATATAC | 1581 |
| Db | 43618 | ACAAAAATATAAATATATTAATAAATAAATAAACAATCTTAATAAATAATTAATTAATTC | 435559 |
| OY | 1582 | ACTATGATTAAGCCATATCAACATATTTATCTTGTCTATCTCTTGTATTTTCAACA | 1641 |
| Db | 43558 | AAAAATCGATTCACACATCAAAACAAATAAATTTCAAAAAACAAATTAATAAATAAATAA | 434999 |
| OY | 1642 | CAATGATATAATTAACGCTCAAAAAATATAGAAATCAACGTTCAACCATATATATTAAGCA | 1701 |
| Db | 43498 | CCAATTTTAAAAATTTACTCTCAACATTAABAAAAAATAAACAATAAATAAATAATATA | 434339 |
| OY | 1702 | GTCAAATACCATGCTTTAATCATCCACTAATGATGTACATATTAATCATTAACAACTTAAG | 1761 |
| Db | 43498 | AAAAATATAAATTTCTTAAACAAAAAACTACCAATTTTAAACAAAAATATATATTTAAAA | 433799 |
| OY | 1762 | TATATCTTAAATTTCTCTAGCTAATTTATATAACAATTCCTAT-CTCTATCTCATATA | 1820 |
| Db | 43378 | CATTAACGAAAAAATAATTCCTTATATATAAATAAATAAATCTACTTACATTAACAA | 433139 |
| OY | 1821 | ACCAATTTGTATCGATTTATTAATAGTCAAAATTAACCTTTTACTTATTTATTAATAAT | 1880 |
| Db | 43318 | ATCAATATACGTAAAAAATAAATAAATACATATCAAAAAATATCTAAATATTTTCTTATTT | 432559 |
| OY | 1881 | ATGTTTCACACCTTTTCACTACATCATGTATATAATGAACATATATTTACATTAATAATTA | 1940 |
| Db | 43258 | CAATATATAAATAAATAATTTTACACACATCCAAAAAATAAACAATTTTACGTATAAAA | 431959 |
| OY | 1941 | CTAATTAATAAAGCAATACATATATTAATTAATCTAATAACACCTATTTTACTACGTAAAT | 2000 |
| Db | 43198 | AAAAATATCAACCTAACTCTAATCTTCTATTTATATATACCAAAAAATTAACAAACACACAT | 431399 |

CC The invention relates to a nucleic acid comprising a sequence of 18
 CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
 CC bisulphite, of genes associated with tumour suppression and oncogenes
 CC having a sequence taken from 536 (actually 533 since numbers 408, 458 and
 CC 500 are missing from the sequence listing) sequences (Ss) and sequences
 CC complementary to (Ss). The nucleic acid may be a peptide nucleic acid-
 CC oligomer (PNA) of at least 9 nucleotides and may form part of a set of
 CC probes for detecting the cytosine methylation state and/or single
 CC nucleotide polymorphisms and also to be used in an array for analysing
 CC diseases associated with CpG dinucleotides e.g. cancers and tumours. The
 CC probes can also be used in a method for ascertaining genetic and/or
 CC epigenetic parameters for the diagnosis and/or therapy of existing
 CC diseases or the predisposition to specific diseases, by analysing
 CC cytosine methylations. The parameters may be compared to another set of
 CC genetic and/or epigenetic parameters, the differences serving as basis
 CC for diagnosis and/or prognosis events which are disadvantageous to
 CC patients. The present sequence is one of the 533 genomic sequences
 CC derived from tumour suppressor genes and oncogenes. Sequences with even
 CC numbered Seg ID numbers are the complementary sequence of the
 CC corresponding odd numbered sequence (e.g. ID 2 and ID1, ID 536 and ID
 CC 535, except for those whose partner sequence is missing). Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 7814 BP; 1677 A; 101 C; 1779 G; 4257 T; 0 U; 0 Other;

Query Match 3 3%; Score 155.2; DB 4; Length 7814;

Best Local Similarity 44.0%; Pred. No. 1,4e-15; Mismatches 1188; Indels 17; Gaps 6;

Matches 945; Conservative 0; Mismatches 1188; Indels 17; Gaps 6;

QY 371 TACATATTAATCAATCAAAACAAATATAGTACTATCTATTAATTAATAAAATCCTC 430
 Db 5641 TACCTATTCATATATTAATAATAAACTAAATATATATACATATTAATAATAATA 5582

QY 431 AATAGCTTTAGCAAAACCTCAAAATATATAGACACTTAATCTATTAATATATTCAA 490
 Db 5581 AATTAACAAACCTAAACCTTACAAAATTAATAATAATAAAACCTAATAATAATAAA 5522

QY 491 ATTATTAATCTATCATCATATATTTAGCATTAATATACATCTGACATGATTAACATAGT 550
 Db 5521 AATTAATCTTAATAAAACACATATACATCTATTTCAATTAATAATAATAAAACCTTAAC 5462

QY 551 AATACATTTCTTACCTAGATATATTAATTAATAATAATAATAATAATAATAATAATA 610
 Db 5461 TTCTCATTAATACATATATTCATTAACAAACCTAATTTATACCCCTTACATTTATAC 5402

QY 611 TTGATTAATATAGTAAATAATGAATATGATTTTGTGACGATCATTTGGCACA 670
 Db 5401 AAAAAATATATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 5342

QY 671 GCATATGCAATTCGCTGTCAGCAATGATATATACATTTATGAGGAGCTGATCAT 730
 Db 5341 ATTAACACACACGATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 5282

QY 731 AGAATATTAATACATATTAACACTTACGCAAAAAATTTAAATATTTATCCGACATATCAT 790
 Db 5281 TAACCTTACACCCCATCAACCAATCAATTAATACTTATCTTTAATAATAATAATAATA 5222

QY 791 CTACGACACACATATATGCAACGACATATATAGACGAATTTATCTGACACAAATACA 850
 Db 5221 AAACAAATAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 5162

QY 851 TGAATTA---TCTTAATCTTTCTTACACAAATTTATGACACATATATGACAAATTAACA 907
 Db 5161 TTAATTAATTTCTTAAAAAACCCTTAATAAATCTTAATAAATCTTTAAAAAATCCCA 5102

QY 908 CACAAACAGCATATGTAATAATAATACTCCATATTAATTTGTAGTAAGATATGCAATTT 967
 Db 5101 ATCAATATCAAAATCAATTAATAATAATAATAATAATAATAATAATAATAATAATAATA 5042

QY 968 ACATCACTCAATTTCCAGTGAATATAGCAAGAAATTTTACATATATATCAATTTT 1027

Db 5041 ATAAATTTCTAATATTTACTATCAATTAATAATAAAACACTTAATTAACCTAAAAATCCACTACA 4982

QY 1028 ATACTCTGTGCTCCAAATTTTGTCTAAAGAAATTCAGACACATCTTCTTGTAGTATAGTA 1087

Db 4981 AAAAAACAAACCTATCTTAATTAATAATAATAATAATAATAATAATAATAATAATAATAATA 4922

QY 1088 CTCTGTGTGATTAATAAGAACTTGGGATCATTTGATTAAGAAACATTAAGTATGATGTT 1147

Db 4921 AAAAAACATTAATTTCAAAATTTCTTCCACACAAAAAAACCTCACAAAAATCCTACAAA 4862

QY 1148 CTAAATAATTAATATCCATCAAGATATATATAGTGTACAGATTGGAGCTGATTAAGAAC 1207

Db 4861 CCAAAAAACACACACACACACACACACACACACACACACACACACACACACACACACACAC 4802

QY 1208 ATAAATTTGCAATTCATGATGTGAATATATGCTGAAAAAATTTAGGTAAATATGCTGTTGCT 1267

Db 4801 ACCCCCAAAAAAATCAAAATCTCTTAATAAAACAAAAATCTTAATAATAATAATAATAATA 4742

QY 1268 ACTGTATTAATCAAGCATATGATGATTAATAAAACATTAATATATAGCAAAAAATCATTTCA 1327

Db 4741 TTCCCAAGCTCACTTAATAAACCCTTAATCTCAATTAATAATAATAATAATAATAATAATA 4682

QY 1328 ATGATCTTCATATACATTAATTTGTCATATGCTTGGAGATCTAATATTAATTAATTAATTA 1387

Db 4681 ATAAAT---TAATAATAACAAATTAATAATAATAATAATAATAATAATAATAATAATAATA 4624

QY 1388 ACAGACATTCACGACATATGCTTTGGACTGAAATAGCAAAAGGTAGAAATATTAAT 1447

Db 4623 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4564

QY 1448 ACATTAATTAATCAACACCTTAAGCTTTGTAAGAAACACGTACTGTAAACACCTGATA 1507

Db 4563 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4505

QY 1508 TCATTAGCAAAAAAATCTTAATGTAGAACTACATTTGCAATCTATTTCAATTTATTA 1567

Db 4504 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4445

QY 1568 CATGAGATATATCATCTAGATTAAGCCATATCAACATTAATCTTATGCTATCATCTT 1627

Db 4444 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4385

QY 1628 TGAATATTTCAACACATGATTAATTAAGCGTCAAAAATTAAGATTAACACGTTCAACAC 1687

Db 4384 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4325

QY 1688 ATATATTAATAAGCATCAATACCCATCTTAACATCACTAAGTATGTAACATATTAAC 1747

Db 4324 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4265

QY 1748 ATAAACAACTTAAGTATTAATTTCTTACGTAATTAATTAATAACAAATTCCTATCT 1807

Db 4264 CAACATATTCATATTAATTAACAAATAATCAACATTAATTCATCAACAAATTAAT 4205

QY 1808 CCTATTCATTAATCAATTTGATCAATTTATCAATTAATTAATTAATTAATTAATTAATTAATTA 1867

Db 4204 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4145

QY 1868 TTTAATAATAATATATGTTGGACACTTTTCACTACATCAATTAATTAATTAATTAATTAATTA 1927

Db 4144 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 4085

QY 1928 TTAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1980

Db 4084 TAGCAAAACACGCAAAACACAAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 4025

QY 1981 ACTATTTTACATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2039

Db 4024 CCGAAATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3965

QY 2040 ACATATTTCAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2099

Db 3964 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3905

of the nucleic acid inhibits proliferation of a cell. Also included are (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid, (2) a host cell containing the vector, (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid, (4) an antibody capable of specifically binding the polypeptide, (5) producing the polypeptide, (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation, (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation, (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts, (9) manufacturing an antibiotic, (10) profiling a compound's activity, (11) a culture comprising strains in which the gene product is overexpressed or underexpressed, (12) determining the extent to which each of the strains is present in a culture or collection of strains, or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acid are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the large-scale prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPAC.

```

SQ Sequence 1065 BP; 358 A; 172 C; 246 G; 287 T; 0 U; 0 Other;
Query Match 3.3%; Score 155; DB 7; Length 1065;

```

| | | | |
|----|-----|----------------------------------------------------------------|-----|
| QY | 584 | AAATAATATTAATTTTGTCTGGTTCCTGTGATAATAATATAGGATAAAAAATATAAATTCGT | 64 |
| Db | 7 | ACTATATATGGAATTATCATATATTATTTTGGAGGAATATCCATGAAACAAAAATCGCA | 66 |
| QY | 644 | ATTTTAGGTGCAGGATCATTTTGGCACACCAATAGAAATTGACCTGTACGACATAGGTATA | 70 |
| Db | 67 | GTTCTTGTCGCCGAGATCTTGGGGCACCTCTTAGACAAAATTTGGCAGAAAAGGGCAT | 122 |
| QY | 704 | TCAGTTAATCTTAGGGGACGTGATCATAGAAATATTACATATATAAACACTTACCGAAA | 76 |
| Db | 127 | GATGTCGGATTTGGGGAAATGTACTCTGACAGATGTATGAAATCTAATTAATCATATCA | 182 |

Db 187 AACACAGATTTTTCCTGATTTCGCATCCCGAGATGATTTCGATATTAAGAAATTG 24
Qy 824 GACGAGATTATTCGACACACATAC -ATGATATTCCTACTATTCCTACACACAT 88
Db 247 GCAGAAAGCAGTGATTAACCCAAATCCTATTATTATTTGTTGCCAACAAGGCATTGGT 300

DQ 942 TAATTGTAGTAAGGATACGAATAATCATCCTCAAAATTTCCCGAGAATAAGCAGAG 100
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 367 AAAGGACTGAACAAGATTCGCATTAACGAATCTCAGAAGTTTGGCGAAGAAATTCCA 422

427 GAA--GAAAAAGACAGGCAATTGTTGATATCTGGTCTTAGCCATGCTGAGGAAGTGG 48

544

1122 TGATAGAAACAATAAGTAATGATGTTCTAAAAATATATACCATCAAGATATTATTAGTGTG 11

```

Db      545 TACAAAACCTGTCATGATGATTTATTTACGATTTATACAAATATATGATGATGGCG 604
Cc      1182 TACAGATTTGAGTGCATTTAAGAACATATTCGATTTGATGATGAAATATGCTGGA 1241
Cc      605 TAGAACTGTGACGACTGAAAATATCTGCAATGATGTCGGCGCTATCCATGATG 664
Cc      1242 AAAATTTAGTAATTAATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1301
Cc      665 TAGGATTCGAGATGATGCAAAAGCAGCATGATGATGATGATGATGATGATGATGATG 724
Cc      1302 CACATATTAATGCAAAAATATTCATGATGATGATGATGATGATGATGATGATGATG 1361
Cc      725 GTCTGTGCTGCGGATGGGT-----GCCATCTCTGACATTTATGCTGATGATG 778
Cc      1362 TTGAGATCTAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATG 1421
Cc      779 TGGGGGACTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 838
Cc      1422 AAATAGAAAAGTAGAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1481
Cc      839 TGTAGGACAGGAGACATTAAGCTGAAAGATTTCTAGAAAACATGGGAAATGCTGTTGAG 898
Cc      1482 GAACGAGTACTGTAAACCACTGATATCATTTAGCAAAAATCTTAATGATGATGATGATG 1541
Cc      899 GTGTGGCAAGCAAAAGCAGCATGATGATGATGATGATGATGATGATGATGATGATG 958
Cc      1542 TTTCATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 1601
Cc      959 TTCTCTCAACATCTATTAATGTTTGTATATGAGAGATATCAAAAAGCAGCAAAAG 1018
Cc      1602 ACATATTAT 1610
Cc      1019 AATCATGT 1027

```

RESULT 21

AD90723
ID ADC90723 standard; DNA; 1071 BP.

XX
AC ADC90723;

XX
DT 01-JAN-2004 (first entry)

XX
DE E. faecium DNA sequence SEQ ID 350.

XX
KM ds: gene; urinary tract infection; bacteraemia; endocarditis; wound;
abdominal-pelvic infection.

XX
OS Enterococcus faecium.

XX
PN US6583275-B1.

XX
PD 24-JUN-2003.

XX
PF 30-JUN-1998; 98US-00107532.

XX
PR 02-JUL-1997; 97US-0051571P.

XX
PR 14-MAY-1998; 98US-0085598P.

XX
PA (GENO-) GENOME THERAPEUTICS CORP.

XX
PI Doucette-Stamm LA, Bush D;

XX
DX WPI, 2003-799836/75.

XX
DR P-PSDB; ADC94377.

XX
XX New isolated nucleic acid derived from Enterococcus faecium encoding an
PT Enterococcus faecium polypeptide useful for detection, prevention and
PT treatment of a pathological condition resulting from a bacterial
infection.

XX
PT Example 1; SEQ ID NO 350; 243bp; English.

XX
XX

XX
Cc The invention relates to an isolated nucleic acid derived from
Cc Enterococcus faecium encoding an Enterococcus faecium polypeptide having
Cc one of 10 fully defined sequences given in the (or comprising 40
Cc sequential nucleotides chosen from any of the nucleic acids, its
Cc complement or sequences hybridizing to it). Also included are a
Cc recombinant vector comprising the nucleic acid operably linked to
Cc transcription regulatory element, a cell comprising the vector and a
Cc single-stranded probe comprising the nucleic acid. The nucleic acids are
Cc chosen from 3654 disclosed sequences encoding 3654 disclosed proteins.
Cc The nucleic acid is useful for diagnosing pathological conditions
Cc resulting from E. faecium bacterial infection (e.g. urinary tract
Cc infection, bacteraemia, endocarditis, wounds and abdominal-pelvic
Cc infection) and for screening drugs such as agonists and antagonists. The
Cc nucleic acid is useful for recombinant production of Candida albicans -
Cc derived peptides or antisense polypeptides. Pharmaceutical compositions
Cc and vaccines containing the nucleic acid are useful for preventing or
Cc treating Enterococcus faecium infections. The present sequence represents
Cc one if the disclosed E. faecium nucleic acids.

XX
SQ Sequence 1071 BP; 362 A; 172 C; 249 G; 288 T; 0 U; 0 Other;

Query Match 3.3%; Score 155; DB 9; Length 1071;
Best Local Similarity 49.0%; Pred. No. 1.4e-15;
Matches 504; Conservative 0; Mismatches 515; Indels 10; Gaps 3;

```

Cc      584 AAATATATATATTTTGTGCTGTTCTTGTATATTAATTAATTAATTAATTAATTAATTAAT 643
Cc      10 ACTATTAATGAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 69
Cc      644 ATTTAGGATGAGATCATTTGGCAGCAGATTAATTAATTAATTAATTAATTAATTAATTA 703
Cc      70 GTCTGTGCTGCTGATTTGGGCGACTGCTTTAGCAAAACATTTGGCAAAAAGCGCAT 129
Cc      704 TCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 763
Cc      130 GATGTCGGTATTTGGGGAATATGTAACCTGAAACAGATTAATTAATTAATTAATTAAT 189
Cc      764 AATTTAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 823
Cc      190 AATCAGATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 249
Cc      824 GACGAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 881
Cc      250 GCAAGACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 309
Cc      882 TAGCAGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 941
Cc      310 TCAATGAGCAGGAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 369
Cc      942 TAATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1001
Cc      370 AAAGGACTAGAAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 429
Cc      1002 AAATTTACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1061
Cc      430 GAA--GAAAAAAGACAGCAATTTGTATTAATTAATTAATTAATTAATTAATTAATTA 487
Cc      1062 CAGACATCTTCTGTTAGATAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1121
Cc      488 CAGTACATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 547
Cc      1122 TGATAGAAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1181
Cc      548 TACAAAACCTGTCATGATGATTTATTTACGATTTATACAAATTAATTAATTAATTAAT 607
Cc      1182 TACAGATTTGAGTGCATTTAAGAACATATTCGATTTGATGATGAAATATGCTGGA 1241
Cc      608 TAGAACTGTGACGACTGAAAATATCTGCAATGATGATGTCGGCGCTATCCATGATG 667
Cc      1242 AAAATTTAGTAATTAATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1301
Cc      668 TAGGATTCGAGATGATGCAAAAGCAGCATGATGATGATGATGATGATGATGATGATGATG 727

```


| | | | |
|-----------|---------------------------------------------------------------------------|----------------------------------------------------------------|------|
| OY | 1302 | CACATATATATAGCAAAAAATCATCTCAATAGATCTTCATACATTAATGGTCCATCATGTC | 1361 |
| OY | 1302 | CACATATATATAGCAAAAAATCATCTCAATAGATCTTCATACATTAATGGTCCATCATGTC | 1361 |
| Db | 728 | GTCCTTGAGTGGCGCATGGGT-----GCCATTCCTGTGACATTTTCGCTTTGAAGTGTC | 781 |
| OY | 1362 | TTGGAGATCTATATTAATTAACATGTACACAGAACTTCAAGCAATATGCTTTTGACTAG | 1421 |
| Db | 782 | TGGGGGACCTTGATCCCTTACTTGTTGACAGAGTTCATCTTCGCAATGGCGGTGCCGTAACT | 841 |
| OY | 1422 | AAATAGCAAAAGGTGAAATATATAATTAATCACTTAATGTATGCACACTTAAGCTGTGAAG | 1481 |
| Db | 842 | TGTTAGGACAGGGAATATAGCTGGAAGAAGTTCTTGAAAACTGGGAAATGGTCGTTGAG | 901 |
| OY | 1482 | GAACCACTACTGTGTAACCACTGATATCATTTAGCAAAAAAACTTAATGTAGAACTACCA | 1541 |
| Db | 902 | GTGTGGCAACGACAAAGACAGAGTAGAGCTCAACAGCTAGAGTGTGAGATGCCAA | 961 |
| OY | 1542 | TTTGATATCTATTATTAACAATTTATATACATGAGAAATATACATTAAGCAATATCA | 1601 |
| Db | 962 | TTACTCAAAACATCTATTAATGTTTGTATTAAGGAAAGATATCAAAAAGACGAAAG | 1021 |
| OY | 1602 | ACATATTAT | 1610 |
| Db | 1022 | AAATCATGT | 1030 |
| RESULT 22 | | | |
| ID | ABL3396/C | | |
| ID | ABL3396 | standard; DNA; 5641 BP. | |
| XX | ABL3396; | | |
| XX | AC | | |
| XX | DT | | |
| XX | 26-MAR-2002 | (first entry) | |
| DE | Human immune system associated gene SEQ ID NO: 1369. | | |
| XX | XX | | |
| KW | Human; immune system disease; cytosine methylation; antiasthmatic; | | |
| KW | antiarteriosclerotic; antiangiogenic; cytosarctic; neoplastic; | | |
| KW | neuroprotective; anti-HIV; anticonvulsant; ophthalmological; | | |
| KW | antirheumatic; antiarthritic; antidiabetic; antiparasitic; | | |
| KW | antitumorigenic; cancer; eye disease; arteriosclerosis; anaemia; | | |
| KW | acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy; | | |
| KW | neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene; | | |
| XX | ds. | | |
| XX | XX | | |
| OS | Homo sapiens. | | |
| XX | XX | | |
| EN | WO200200928-A2. | | |
| XX | XX | | |
| PD | 03-JAN-2002. | | |
| XX | XX | | |
| PF | 02-JUL-2001; 2001WO-EP007537. | | |
| XX | XX | | |
| PR | 30-JUN-2000; 2000DE-01032529. | | |
| PR | 01-SEP-2000; 2000DE-01043826. | | |
| XX | XX | | |
| PA | (EPIG-) EPIGENOMICS AG. | | |
| XX | XX | | |
| PI | Olek A, Piepenbrock C, Berlin K; | | |
| XX | XX | | |
| DR | WPI; 2002-130909/17. | | |
| XX | XX | | |
| PT | Nucleic acid comprising fragment of chemically modified gene, useful for | | |
| XX | diagnosis and treatment of diseases associated with abnormal cytosine | | |
| XX | methylation. | | |
| XX | XX | | |
| PS | Claim 1; SEQ ID NO 1369; 32bp + Sequence Listing; German. | | |
| CC | The present invention provides a number of human immune system associated | | |
| CC | genes which are modified by the methylation of cytosines. The sequences | | |
| CC | can be used in the diagnosis and treatment of immune system disorders, | | |
| CC | including eye diseases such as retinopathy, neovascular glaucoma and | | |
| CC | macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid | | |

| Query Match | 3.3% | Score 153 | DB 6 | Length 5641 |
|-----------------------|-------------------------------------------------------------------------|-------------------|-----------|-------------|
| Best Local Similarity | 43.9% | Pred. No. 1.5e-15 | | |
| Matches 1027 | Conservative 0 | Mismatches 1280 | Indels 34 | Gaps 7 |
| XX | Sequence 5641 BP, 1442 A, 34 C, 930 G, 3235 T, 0 U, 0 Other: | | | |
| CC | leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis, | | | |
| CC | rheumatoid arthritis, psoriasis and inflammatory ulcerative bowel | | | |
| CC | diseases. The present sequence is a gene of the invention | | | |
| CC | | | | |
| XX | | | | |
| QY | 176 ATTTCCTCAATCTCCCAATTTTAAACAGCAGCAAGATATTTAAGTAACAGAACACCTG 235 | | | |
| DB | 4946 ATTAAATPAAAACCTAAAATTAACCAATACATTTTCTTAAATTTAAAATCTTA 4887 | | | |
| QY | 236 TACTCTTATGCAAGACCTATATAAAGCTATCAAGCTCTTACATTAAGAAACAAGTAT 295 | | | |
| DB | 4886 CACCCAAATTTATATATTTTAAAAATAAAAAATA---TAACAAAATTTTAT 4830 | | | |
| QY | 296 TGAAGAAGAACCTCTCTCTGTGTCATCCATATTTCCATATTTTAAATATCATAGA 355 | | | |
| DB | 4829 ACGATATATATTTATTCACCTACCTCAAACTAAAAAAGCTCATATTTAAAAATATATA 4770 | | | |
| QY | 356 GTTCTTTATTCGACATACATATTTAAATCAATCAACAAATATATAGTATCTATTT 415 | | | |
| DB | 4769 TAATCTTATTTTCAATCCATACATCTTACACATCTTCTTAAATATTAATTAATAT 4710 | | | |
| QY | 416 AATAAAAAATCCCATAGT---CTTAGCAAGACCTCAATATATAGACACTTA 471 | | | |
| DB | 4709 ATTATATATATATATATATTTAAATATATATATATATATATATATATATATATTA 4650 | | | |
| QY | 472 ACTATTTATATATATTTCAAAATTTTATATACCTACATCATATTTAGATATATACATC 531 | | | |
| DB | 4649 AAAAAAATTTAAATATTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATATA 4590 | | | |
| QY | 532 TGACATATTTAGACATGTAGACATTTCTTAAACAGTAGAATATATCAATTAATATAT 591 | | | |
| DB | 4589 ATATCTATCTAAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAAT 4530 | | | |
| QY | 592 ATAAATTTTGAAGTGTCTGTGATATATATTAATTAATTAATTAATTAATTAATTAATTA 651 | | | |
| DB | 4539 ATACTTTAAATCAATCTTTCCCTACATTAATTAATTAATTAATTAATTAATTAATTAATTA 4470 | | | |
| QY | 652 TGCAGATTCATTTGGCAGAGATAGCAATTTGCACGTGCGACAGTGTATATCAGTTAA 711 | | | |
| DB | 4469 TAAATCTCAATTAATTTCCAAATTTCTTAAATCTCAATTTCTTCAATAGATCATATA 4410 | | | |
| QY | 712 CTTATGGGAGCGTGATCATAGAAATATTACACATATTAACACCTTACCGAAAAATTTAAA 771 | | | |
| DB | 4409 ATATAT-----AAATTAATAAAAAATTAATTAATTAATTAATTAATTAATTAATTAAT 4357 | | | |
| QY | 772 ATATTACCCACATATCATCTACAGCAACATATATGCAACGCAATATAGCGAGT 831 | | | |
| DB | 4356 AAACATTTAACCTCAAAAATTTATCTATTTATTTATTTTAAACGAAATTCAAAAATA 4297 | | | |
| QY | 832 ATATATGACACATATCATATGATTTATCTTAACTATTCTTACACAAATTTACGACCAT 891 | | | |
| DB | 4296 ATATTAATTAATTAATAATCTATCTCATCCCATATCAACAAATTAATATATACCTCAAT 4237 | | | |
| QY | 892 ATGTACCAATATCAACAAACAGCATATGTGTAAATAATATCCTCATATTAATTTGAG 951 | | | |
| DB | 4236 ATTAACAAATTTCAAGAAACCTTCAATTTCTCCAAATTAATAATAATAATAATAATA 4177 | | | |
| QY | 952 TAAAGTATGAAATTAACATCACTCAATTTCCAGTGAATAGCAGAGAAATTTTACA 1011 | | | |
| DB | 4176 TAAATTTATTAACATATCAAAATTTCTTAATATATAAAAAATATACACAAATCTATTAAT 4117 | | | |
| QY | 1012 ATATATTCCAATTTTATATCTCTGTGTCAGATTTGTCAAGAAATTTGACAGACATCT 1071 | | | |
| DB | 4116 CTTTAAATTTATTAATAATTTTCTAATATCAAAATTTATTAACAAATTTTAAATAATA 4057 | | | |
| QY | 1072 TCCTTTAGTAGTAGTACTTGCTGTGTATTAAGAACTTGGTGATCATTTGATGAAC 1131 | | | |
| DB | 4056 AAATTAATCTAAATTTTATCTTAATATTAACATATTAATTAATTAATTAATTAATTAATTA 3957 | | | |

QY 1132 AATAAGTAAAGTCTTAAAAATATATACATCAAGATATATAGTGTACAGATTGG 1191
 Db 3996 AAAAAATTAACAAATATATCTTACACATCAAAATATTTAAAAATCTTAAAAACATATATATA 3937
 QY 1192 AGCTGATTAAGAAATATATGCAATTCGATGTGGAATATCGCTGGAAAAATTTAGG 1251
 Db 3936 TAATATTAATAATTAACAAATCAATTAATAAACTTCACAAAATTAACAAAAACGTAAATTT 3877
 QY 1252 TAATATGCTGTGCTACTGTTATAAATAAGCAATGATGAATTAATAACCTATATAT 1311
 Db 3876 AACAAATTAATTAATTAATACATTAATTAATTAACCAATTAATTTCTAATAATACATCTAT 3817
 QY 1312 AGCAAAAAATCATCTAATGATCTTCAATTAATTAATGCTCATGATGTCTTGAGATCT 1371
 Db 3816 ATTAACAAATTAATAACCTATATTCAAAAATTAACACATTA-----ATAACT 3767
 QY 1372 AATATTAACATGTAACAGAAATTCACGAATATGGCTTTGAGACTAGAAATAGAAA 1431
 Db 3766 ATATATTAATTTGAAAAACATTAATTCGAACAAATTTTATATCTTATCAAAACATTTTA 3707
 QY 1432 AGGTAGAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1491
 Db 3706 AAACATACCTAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3651
 QY 1492 TGTAAACACATGATATCAATTAAGCAAAATTAATTAATTAATTAATTAATTAATTAATTA 1551
 Db 3650 TATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3591
 QY 1552 TATTAACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1611
 Db 3590 AATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3531
 QY 1612 TAGTCTATCACTCTTGTGTAATTCACACAAATTAATTAATTAATTAATTAATTAATTAATTA 1671
 Db 3530 TCTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3471
 QY 1672 ATACAAAGTTCACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1727
 Db 3470 AATAACCAACCCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3411
 QY 1728 CTAAAGTATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1787
 Db 3410 TATTCATATTCACCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3351
 QY 1788 AATAAACAATTCCTATCTCTATCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1847
 Db 3350 AAAAAAATTAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 3291
 QY 1848 TCATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1907
 Db 3290 AAAAACTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3231
 QY 1908 TGTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1967
 Db 3230 AAAAAAATTAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 3171
 QY 1968 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2025
 Db 3170 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3111
 QY 2026 TCTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2085
 Db 3110 TTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3051
 QY 2086 TAGTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2145
 Db 3050 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2991
 QY 2146 ATGAAGAAGATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2205
 Db 2990 CTCAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2931

QY 2206 CTAAACATAGGAACCATAGTATATATATATATATATATATATATATATATATATATATATATAT 2265
 Db 2930 AAAAAAATTAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 2871
 QY 2266 AATAATTAAGAAATATATGTTATATATATATATATATATATATATATATATATATATATATATAT 2325
 Db 2870 ATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2811
 QY 2326 CAAAACTCAAGTCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2385
 Db 2810 AAAAAATTAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 2751
 QY 2386 TATCAAAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2445
 Db 2750 TCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2691
 QY 2446 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2505
 Db 2690 AACAACTTCTCAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 2631
 QY 2506 A 2506
 Db 2630 A 2630

RESULT 23
 AAV74824/C
 ID AAV74824 standard; DNA; 918 BP.
 AC
 XX AAV74824;
 DT 16-MAR-1999 (first entry)
 XX
 DE Staphylococcus aureus contig SEQ ID #513.
 XX
 KW Computer readable medium; vaccine; S. aureus infection; immunodetection;
 KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
 KW skin infection; surgical wound infection; scalded skin syndrome;
 KW toxic shock syndrome; ds.
 XX
 OS Staphylococcus aureus.
 XX
 FH Location/Qualifiers
 FH Key 781..840
 FH misc_feature
 FT /tag= a
 FT /note= "these bases represent a line of missing text in
 FT the sequence listing in the specification. They are
 FT included to maintain the nucleotide numbering given in
 FT the specification for this DNA sequence"
 FT
 XX
 EN EP786519-A2.
 XX
 PD 30-JUL-1997.
 XX
 PF 07-JAN-1997; 97BP-00100117.
 XX
 PR 05-JAN-1996; 96US-0009861P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Kunsch CA, Choi GH, Barash SC, Dillon PJ, Fannon MR, Rosen CA;
 DR WPI; 1997-374922/35.
 XX
 PT Polynucleotide(s) and proteins derived from Staphylococcus aureus -
 PT stored on computer readable medium and used in the production of anti-
 PT S. aureus vaccines.
 XX
 PS Claim 1; Page 1440-1441; 3271pp; English.
 XX
 CC This sequence represents one of 5191 Staphylococcus aureus DNA sequences
 CC of the invention. The DNA sequences are recorded on a computer readable
 CC medium, preferably selected from a floppy or hard disk, random access

XX 26-MAR-2002; 2002MO-EP003401.
 XX 26-MAR-2001; 2001US-0278333P.
 XX (EPIC-) EPIGENOMICS AG.
 XX Berlin K, Braun A, Disler J, Gueig D, Howe A, Mueller J,
 XX Olex A, Piepdrock C, Adorjan P, Grabs G, Lesche R, Ley E,
 XX Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T, Pelet C,
 XX Schwope I, Ziebach H;
 XX WPI; 2003-018942/01.
 XX
 PT Detecting and differentiating between hematopoietic cell proliferative
 PT disorders, comprises contacting a target nucleic acid with a reagent that
 PT distinguishes between methylated and non-methylated CpG dinucleotides.
 XX
 PS Claim 28; SEQ ID NO 240; 117P; English.
 XX
 CC The present invention describes a method for detecting and
 CC differentiating between hematopoietic cell proliferative disorders
 CC associated with at least 1 gene and/or their regulatory regions in a
 CC subject. The method comprises contacting a target nucleic acid in a
 CC biological sample obtained from the subject with at least 1 reagent,
 CC which distinguishes between methylated and non-methylated CpG
 CC dinucleotides within the target nucleic acid. AB203861 to AB211118
 CC represent specifically claimed nucleotide sequences from the present
 CC invention. Oligonucleotides from the present invention can be used: for
 CC differentiating between healthy hematopoietic cells and proliferative
 CC disorder hematopoietic cells; for differentiating between acute
 CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
 CC determining the cytosine methylation state and/or single nucleotide
 CC polymorphisms (SNPs) of hematopoietic cell proliferation disorder
 CC related sequences and their complements; and as primers for the
 CC amplification of hematopoietic cell proliferation disorder related
 CC sequences. The nucleotide sequences from the present invention can also
 CC be used for detecting a predisposition to, differentiation between
 CC subclasses, diagnosis, prognosis, treatment and/or monitoring of
 CC hematopoietic cell proliferative disorders. The present method enables a
 CC highly specific classification of hematopoietic cell proliferative
 CC disorders allowing for improved and informed treatment of patients
 XX
 SQ Sequence 8056 BP; 371 A; 371 C; 371 G; 3603 T; 0 U; 0 Other;
 Query Match 3.2%; Score 150.6; DB 7; Length 8056;
 Best Local Similarity 44.5%; Pred. No. 7; 7e-15;
 Matches 1097; Conservative 0; Mismatches 1324; Indels 42; Gaps 11;

OY 408 ATCTAATTATTAATAAATTCCTCAATGCTTTAGCAGACTACCAATTAATGAGACA 467
 DB 795 ATATTTAAATTAATAATTAATTAACGATTAATTAATAATTTATTAATAAATTCGTTT 854
 OY 468 CTTACTTATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 524
 DB 855 TTTTATTTTATTTTACGTATTAATAAATTCGTAATTAATTAATTAATTAATTAATTAAT 914
 OY 525 TCACATGTGACATGATTAACAAATGTAACCAATTTCTTAACCTAGATTAATACAAATA 564
 DB 915 TCGCGTTTCGTAAATTAATTTTAAATAAATTTTATTTTATTTTAAATTAATTAATTA 973
 OY 585 ATATATATATATTTTACTGTTTCTGATATTAATTAATTAATTAATTAATTAATTAATTA 644
 DB 974 ATTAATAATTAATTAATTAATTTTAAATAATTAATTTTAAATAATTTTATTTTATTTT 1033
 OY 645 TTTTAGGTGAGATCATTTTGACACAGCAATGCAATGCACTGTCGACATGATATAT 704
 DB 1034 TTAATTCGAAATTTATTAATTAATTAATTAATAAATAAATTCGTTT-----TAATTCGTTT 1088
 OY 705 CAGTTACTATATGGGCGGATCATAGAAATATTAACATATTAACATTAACCGAAAA 764
 DB 1089 TTTTAAACGTTTAAACGTTATATATGTTTATTAATTTTTCGATATTCGAAAA 1148
 OY 765 ATTTAAATATTTAACCAATATCATCTACAGACAAATATATGCAACCGCAATATAG 824
 DB 1149 TTTAAATTAAT-----ATATTCGATATTTTAAATTTTAAATTTTAAATTAATTA 1203
 OY 825 ACGAGATATATCGAACAAATACATATATATTAATTAATTAATTAATTAATTAATTAAT 884
 DB 1204 TTTAATTTAATTTTAAATTAATTAATTAATTTTTCGAAATTAATAAATTCGTTAGTA 1263
 OY 885 GCACCATATGACAAATTAACAAACAAACAGCATATGTTAAATTAATTAATTAATTA 944
 DB 1264 TTAATTAATTAATTAATAAAGTAATATTTTAAATTAATTAATTAATTAATTAATTA 1323
 OY 945 TTGTAGTAAGATGATGAAATTAATCACTCAATTTCCAGTGAATAGCAAGAA 1004
 DB 1324 TTAATTAATTTTAAATAAATAATTAATTAATTTTAAATTTTAAATAAATTTT 1383
 OY 1005 TTTTACATATTAAT-CCAAATTTTAACTCTGCGCAAGTTTCTAAAGAAATTTGA 1063
 DB 1384 TTTAATAAATAATATGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1443
 OY 1064 GAACATCTTCTGTAGATATGATCTGCGGATTAATTAATAAAGTGTGATCAT 1123
 DB 1444 ATAAATAATTTTATTTTAAATAATTAATTAATTAATTAATTAATTAATTAATTAAT 1503
 OY 1124 ATGAATAATTAAGTAATGATGTTCTAAATAATTAATTAATTAATTAATTAATTAATTA 1183
 DB 1504 ATAAATAATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTT 1556
 OY 1184 CAGATTGAGCTGATTAAGAAACATTAATGCAATTTGATGGAATTAATCGCTGAAAA 1243
 DB 1557 ---TAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTA 1613
 OY 1244 AATTAGATATATGCTGTGCTGCTGATTAATTAATTAATTAATTAATTAATTAATTA 1303
 DB 1614 AAAAATGAAATTAAGCAATCGTAATTAATTAATTAATTAATTAATTAATTAATTA 1673
 OY 1304 CTATATATAGCAAAAAATCAATGATGATCTTCATTAATTAATTTGGTCATATGCTT 1363
 DB 1674 TTTTATTTTATTAATAAATTAATAAATTAATTAATTAATTAATTAATTAATTTT 1733
 OY 1364 GGAATCTAATTAATTAACATGATACAGCAATTCACGATATGCTTTGAGCTAGAA 1423
 DB 1734 AATTATTTAAATAATTAATAAATTAATTAATTTTAAATTAATAAATAATTAATTA 1793
 OY 1424 ATAGAAAGGTAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1483
 DB 1794 GTTATTTATTTTATTAATTAATTAATTAATTAATTAATTTGTTTAAAAA-----AAAAA 1846
 OY 1484 ACCAGTACTGTAAACACATGATATCATTTGCAAAAAAATTAATTAATTAATTAATTA 1543

| | | | |
|----|------|---------------------------------------------------------------|------|
| Db | 1847 | AAAAAAAAAAAAAAAAATTAATTAAATTATTTTATTTTATTTATTAATAAATAAATAA | 1900 |
| Qy | 1544 | TGCATATCGATTTACATTTATTAATGAGATTA-TCACATGATTAAGCATATCCAA | 1602 |
| Db | 1907 | ATTTTATTAATAAATAAATAAATAAATTAAATTAATTAATAAATAAATAAATAAATTTTAA | 1966 |
| Qy | 1603 | CATATTAATCTTAGTCATATCCCTTGATTAATTCACACATGATTAATTAACGCTCA | 1667 |
| Db | 1967 | AATAAATAATTAATTAATTTTATTTAATAAATTTAATAAATTTAATTAATTTAA | 2026 |
| Qy | 1663 | AAATTAAGATGCAACGTCACCATATTAATAAGATCAATACCCATCGTTACA | 1722 |
| Db | 2027 | TAAATTAATTAATTAATTAATTTTATTTTAAATAATTTATTAATAAATTTTAAATA | 2086 |
| Qy | 1723 | TCCAACTAAGTAGTACATATTAATTAACAACCTAAGTATTAATCTAATAATTCCTAGC | 1782 |
| Db | 2087 | TTACAAATTAATA-ATTATTAATAATTTATTAATTAATTAATAAATTTTATTTATTA | 2145 |
| Qy | 1783 | TAAATTAATTAACAATTCCTATCCCATCTCAATTAACCAATTTGTTATCAGATTATA | 1842 |
| Db | 2146 | AATTAATAATTAATAAATAAATAAATAATTAATTAATTAATAAATAAATAAATAA | 2209 |
| Qy | 1843 | ATAAGCTAAATTAACCTTTACTATTAATTAATAAATAATGATTGCACTTTTCACTA | 1902 |
| Db | 2206 | AAAAATTAATAAATAAATAAATAATTAATTTAATTAATAAATTTTATTAATTTAATA | 2265 |
| Qy | 1903 | CATCATGATTAATGAACATTAATTTACATTAATACTAATAAATAAAGAACCATACAT | 1962 |
| Db | 2266 | AATTAATAAATAATTAATTTTATTAATTAATTAATTTAATTAATAATTAATTAATTTT | 2325 |
| Qy | 1963 | ATATTAATTAATTAACAACCATTAATTAATTAATTAATTAATTAATTAATTAATTAAT | 2022 |
| Db | 2326 | ATTTAATAAATAATTAATTTAATTTAATAATAATTTTATTTTATTAATAATTAATTA | 2385 |
| Qy | 2023 | AAATTCATCATCCCTGACATTTTCAACAATAAATGATTAACGT-----TTTTT | 2076 |
| Db | 2386 | AAATTTATTTTATTTTGCATAATTAATAAATAAATAATTTATTTTAAAAAATTAAT | 2445 |
| Qy | 2077 | TATGCCATTAAGTCTTAATTAATAATGCTAATTAACCAATAAGGTTAAATATACATTC | 2136 |
| Db | 2446 | TTTTTTTATTTTATTAATTAATTAATTTTATTTATTTGTTTAAATATTAATAATTA | 2505 |
| Qy | 2137 | TGCAACAATGAAGAGATACATACGCGCAATAATTAACCAATTTCTAATATA | 2186 |
| Db | 2506 | ACGAAAAATTAATAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTA | 2565 |
| Qy | 2197 | CCACTACTTAACATGAGAACCAATGATATCATATCCTAAGTAGATATCCCTCAG | 2256 |
| Db | 2566 | ATTAATAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAATTA | 2625 |
| Qy | 2257 | CATGATCAATTAATAAGAAAAATCGTATGATTTTATTAACATTAACAATAATTCGA | 2316 |
| Db | 2626 | AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA | 2685 |
| Qy | 2317 | CTTCACAACAAACTCAAGTCTATATTAACAATCAAGATTAATATTAATCTCACC | 2376 |
| Db | 2686 | TTTTTTTAAATTTATTAATTTTATTTTAAATAAATAATTCGTTAAATTAATTA | 2745 |
| Qy | 2377 | AACCATATGTTATCAAAATTCATTAATACATATATCCCTTTAAATAAGATGACTCACA | 2436 |
| Db | 2746 | TAAATAATAAATAAATAATTAATGTTATTTTATTAATAATTAACAATTCGTTAT | 2805 |
| Qy | 2437 | TTCAATACATAATAATTAATACATCAATGTTTATTCGAAAAATTAACAATCCAGAT | 2496 |
| Db | 2806 | TAAATATGTTTATTAATTAATAACGATATTTTATTAATAAATAAATAATTCGAATTA | 2865 |
| Qy | 2497 | TTT 2499 | |
| Db | 2866 | TTT 2866 | |

[illegible]

Db 5099 TTTCAAAAAAATAACAAAAAACCCTCTTAAAAATTAACCTTAACTACACT 5040
Qy 470 TAACTATTATATATTTCAAAATTTATTAAGTACTCATCATTTAGCATATATCACA 529
Db 5039 AAA--AAATAAAAAATAACCAATTAACGACCCAAACAAATTTCTTCATTCATA 4983
Qy 530 TCTGACATGATTAACAATAGTAACATTTCTTAACATAGATATATATATATATAT 589
Db 4982 CCTAAATCTTTCTTAATACAAACAAAAAACAATTAACCTTACCTTAAAAAA 4923
Qy 590 ATATATATTTTGACTGGTTCTTGATATATATATAGTAAAAAATGAAATCAGATTTTA 649
Db 4922 TTTAATCTCTCTCTACGCTATAAAT--AAATACCAAAAAAACAATCTTAATATAT 4866
Qy 650 GGTGACGATCTTTGGACAGCAATAGCAATTCGACTGTGACCATGTATATCAGTT 709
Db 4865 AAAACCTACAAAAAATACTATATTTCTAAAAAATAATCAAAATTAACCTTAAAAAA 4806
Qy 710 AACTATGGGACGTGATCATAGAAATTTACACATATTAACACTTACCGAAAAATTTA 769
Db 4805 TAAATATATTTTAAACACAAAAAATAATATAAAAAAAGAAAAAATATATATC 4746
Qy 770 AATATATTAACCATATATCTACGACAAACATATATGACACGCAATATAGACGA 829
Db 4745 TATTAATTAATAAAAAAATAAAAAAATACTATCATCTTAACTTAACTTAAAAAT 4686
Qy 830 GTATATCTGACACATATCATGTATTTATCTTAATCTTACACACAAATTAAGGACC 889
Db 4685 ATTAATAAAAAAATAACATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4626
Qy 890 AATGTACCAATTAACAACAACAACATATGTGTAAAAATCTCCAAATTAATTTGT 949
Db 4625 TTTACTCTATTAATAAAAAAATAAAAAAATAATTAATTAATTAATTAATTAATTA 4586
Qy 950 AGTAAAGTATGGAATTAATTAATCACTCAATTTCCCATGAAATAGCAAGAAATTTTA 1009
Db 4565 AATTAATTAATTAATAAAAAAATAATTAATTAATTAATTAATTAATTAATTAATTA 4506
Qy 1010 CAATATTAATCAATTTTATTAATCTCTGTGTCAGATTTGCTAAAGAAATTCAGAACT 1069
Db 4505 TAAAAATCAAAAACTATTTTACCAAAATTAATAATTAATAAACAACAACTCAAACTAA 4446
Qy 1070 CTTCCTGATGATTAAGTACTGCTGTGTGATTAATAAGAACTGTGGAATCTTATGA 1129
Db 4445 AATTTATTAATCTTTATTAATTTTAAATTAATTAATTAATTAATTAATTAATTAAT 4386
Qy 1130 ACAATTAATTAATGAT-GTTCTAAAAAATAATTAATTAATTAATTAATTAATTAATTA 1188
Db 4385 ATTAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 4326
Qy 1189 TGGAGCTGATTAAGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1248
Db 4335 AATTAATCTCAAAAAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4266
Qy 1249 AGGTATTAATGCTGTGCTACTGTATTAATTAATTAATTAATTAATTAATTAATTA 1308
Db 4265 AAAAAATTAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4206
Qy 1309 TATACCAAAAAATCTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1368
Db 4205 AAAAAATTAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4146
Qy 1369 TCTTAATTAATCAATGT--ACAACAGAACTTCAAGCAATATGCTTTGCTGAATA 1426
Db 4145 TCTTAATTAATAAAAAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4086
Qy 1427 GGAATAGGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1486
Db 4085 CGAATTAATCTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4026
Qy 1487 AGTACTGTAAACCACTGATATCAATTAAGCAAAAAAATCTTAA--TGTAGAACTTCAAT 1543
Db 4025 AATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3966

Qy 1544 TGCATATCTATTTTACATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1603
Db 3965 TAAAAATTAATAATTAATAAAAAAATAAATTAATTAATTAATTAATTAATTAATTAAT 3906
Qy 1604 ATATTAATCTTATGCTATCATCTTTGATTAATTTTCAACATATTAATTAATTAATTA 1663
Db 3905 ATAAATTAACCAATTAACATTAATAAAAAAATAAATAAATAAATAAATAAATAAATA 3846
Qy 1664 AATATGAATATACAGCTTCAACATTAATTAATTAATTAATTAATTAATTAATTAATTA 1723
Db 3845 AAAAAAATAATTAACAGAAAAAATAATTAATTAATTAATTAATTAATTAATTAATTA 3788
Qy 1724 CCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1783
Db 3787 CAATCAAAACAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 3728
Qy 1784 AATTAATTAACCAATTCCTATCTCTATCTCAATTAACCAATTTGATCAGATTAATA 1843
Db 3727 AATCAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 3673
Qy 1844 TAACTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1903
Db 3672 TAACTTCA--AAATTTCAATTAATAAATAAATAAATAAATAAATAAATAAATAAATA 3616
Qy 1904 ATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1963
Db 3615 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 3556
Qy 3555 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 3496
Qy 2024 AATTC--TATCATCTGATTAAT--TTCAACATTAATAATTAATTAATTAATTAATTA 2081
Db 3495 AATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3436
Qy 2082 CCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2141
Db 3435 CAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3376
Qy 2142 CAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2201
Db 3375 TAACTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3316
Qy 2202 TACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2261
Db 3315 CAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3256
Qy 2262 ATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2321
Db 3255 TACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3197
Qy 2322 ACACAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2381
Db 3196 ACCGAAAAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3137
Qy 2382 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2441
Db 3136 GACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3077
Qy 2442 ACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2501
Db 3076 CTAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3017
Qy 2502 ATTTAGGTGAT 2515
Db 3016 AACTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3003

RESULT 27
ACA18468
ID ACA18468 standard; DNA; 1020 BP.

XX ACA18468;
 AC 19-JUN-2003 (first entry)
 DT Prokaryotic essential gene #125.
 DE
 XX Antisense; ds; prokaryotic essential gene; cell proliferation;
 KM drug design; gene.
 XX Enterococcus faecalis.
 OS
 XX W020027183-A2.
 PN
 XX 03-OCT-2002.
 PD
 XX 21-MAR-2002; 2002MO-US009.07.
 PF
 XX 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948893.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 XX (ELIT-) ELITRA PHARM INC.
 PA
 XX Wang L, Zamudio C, Malone C, Haeelbeck R, Ohlsen KI, Zyskind JW,
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH,
 PI WPI; 2003-029926/02.
 DR P-PSDB; ABU14598.
 DR
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 PT
 XX Claim 14; SEQ ID NO 6338; 1766bp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence data for this patent
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

Query Match

3.1%; Score 148.4; DB 7; Length 1020;

Best Local Similarity 49.1%; Pred. No. 1.6e-14;
 Matches 483; Conservative 0; Mismatches 491; Indels 10; Gaps 3;

| | | | |
|----|------|----------------------------------------------------------|------|
| Qy | 629 | AAATGAAATACGATTTTGTGACGAGATCATTTGGCAGACGAAATGCAATTCAGCTG | 688 |
| Db | 4 | AAACAAAAGTCGCTGTTTGTGTCAGGTCCTGGGGGACAGCTTTAGCTCAAGTATTA | 63 |
| Qy | 689 | TCAGCAGATGATATATGATTAAGTATGAGGAGATGATCAAGAAATATATACATATA | 748 |
| Db | 64 | GCCGAAATGACACGAGGCTGTATTTGGGGAATTAACGACGCAATGATGAAATA | 123 |
| Qy | 749 | AACACTTCGCAAAAATTTAAATATTACCCCAATTCATCTACGACACATATAT | 808 |
| Db | 124 | AATACAAAACACACCAACCAATTTATCCAGAAATTAATCTTACCTTCAATTCAG | 183 |
| Qy | 809 | GCAACGACATATATAGACGAGATTTATCTGACAAATAC--ATGATATATCTTAATA | 866 |
| Db | 184 | GCAACGACTGATTTAGCAGCTGCTTATGATGATGATGATGATGATGATGATGATG | 243 |
| Qy | 867 | TTCTTACCAACAATATGACGACATATGATACAAATATACACAAACAGATATGTA | 926 |
| Db | 244 | ACTTAAGCCATTCCTGATGATGACCAAGAGTGGCCCAACATTTAAACCAACAT | 303 |
| Qy | 927 | AAATATCTCAATATTTATTTTGTAGTAAGGTATGCAATATGATCACTCAATTC | 986 |
| Db | 304 | ATTATTCATGCAATGATTAAGGCTGAAACAGATACATTAAGCGGATTTCAAGTATA | 363 |
| Qy | 987 | GTGAATATGCAAGAAATTTTACATATATATCCATTTTATCTCTGTCGCAATT | 1046 |
| Db | 364 | GCGGAGAGATTCAGACAGA--AAACGGCAAGGATTTGATTTATCTGCTCAAGCC | 421 |
| Qy | 1047 | TTGCTAAAGAAATGCAAGAAATCTTCTGATATATGATATGATATGATATGATAT | 1106 |
| Db | 422 | ATGCAAGAAAGTGTGCTTTATATACATATACATATACATATACATATACATAT | 481 |
| Qy | 1107 | AACCTGTGATCATTTGATATGAAACAAATATGATATGATATGATATGATATGAT | 1166 |
| Db | 482 | CAGAGCTGTATGATGCAAGATTTTATGATGATATCTTATGATATATATGATAT | 541 |
| Qy | 1167 | AAGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG | 1226 |
| Db | 542 | ATGATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG | 601 |
| Qy | 1227 | GAATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG | 1286 |
| Db | 602 | GAGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG | 661 |
| Qy | 1287 | TGAATGAAATTAACACTATATATACCAAAATATCTCAATGATCTTATATATTA | 1346 |
| Db | 662 | TAGCAGAAATTAAGCTTTTGGGGTTCGATGGG-----AGCAACCATTAACATTA | 715 |
| Qy | 1347 | TTGCTTCATCATGCTTCTGAGATCTATATATTAATATATATATATATATATAT | 1406 |
| Db | 716 | TTGCTTTAAGGCTTTGGGGATTTATTTTCTTATATATATATATATATATAT | 775 |
| Qy | 1407 | TGCTTTTGAATGAAATTAAGAAAGTGAATTAATATATATATATATATATATAT | 1466 |
| Db | 776 | GCGCTGGGGGATTTTATTTGAAAGGCAATGATGATGATGATGATGATGATGAT | 835 |
| Qy | 1467 | TAAAGCTTGTGAAAGAACAGTACTATAAACCACTGATATATATATATATATAT | 1526 |
| Db | 836 | GCAATGATTTGGAAGGCTTTCAACAACTTAAGAGGCTTATGATTTGCAACAACT | 895 |
| Qy | 1527 | ATGTAAGACTACCAATTTGATATATATATATATATATATATATATATATATAT | 1586 |
| Db | 896 | AAGTCGAATGCAATACAGAAACATATATATATATATATATATATATATATAT | 955 |
| Qy | 1587 | ATAAGGCTATATCAATATATATATATATATATATATATATATATATATATAT | 1610 |
| Db | 956 | AACAAGCAGCAAAAGAAATCATGT 979 | |

RESULT 28

AA55118
 ID AA55118 standard; DNA; 1023 BP.
 AC AA55118;
 DT 13-FEB-2002 (first entry)
 DE Enterococcus faecalis DNA for cellular proliferation protein #546.
 XX
 XX Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;
 XX antibacterial; drug design.
 XX
 OS Enterococcus faecalis.
 XX
 PN W0200170955-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001WO-US009180.
 XX
 PR 21-MAR-2000; 2000US-0191078P.
 PR 23-MAY-2000; 2000US-0206848P.
 PR 26-MAY-2000; 2000US-0207727P.
 PR 23-OCT-2000; 2000US-0242578P.
 PR 27-NOV-2000; 2000US-0253625P.
 PR 22-DEC-2000; 2000US-0257931P.
 PR 16-FEB-2001; 2001US-0269308P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Heselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ,
 PI Yamamoto RT, Xu HH;
 DR WPI; 2001-611495/70.
 XX
 DR P-SDB; AAU35259.
 XX
 PT New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids.
 XX
 PS Claim 27; SEQ ID NO 6755; 511bp; English.
 XX
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the genes,
 CC their use in the discovery of novel antibiotics, the essential genes,
 CC themselves and the encoded proteins. The prokaryotes used are Escherichia
 CC coli, Streptococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
 CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
 CC useful for the identification of potential new targets for antibiotic
 CC development. The antisense nucleic acids can also be used to identify
 CC proteins used in proliferation, to express these proteins, and to obtain
 CC antibodies capable of binding to the expressed proteins. The proteins can
 CC be used to screen compounds in rational drug discovery programmes. The
 CC antisense nucleic acid sequence is also useful to screen for homologous
 CC nucleic acids which are required for cell proliferation in a wide variety
 CC of organisms. The present sequence encodes an essential prokaryotic
 CC cellular proliferation protein. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pat_sequences
 XX
 SQ Sequence 1023 BP; 344 A; 172 C; 237 G; 270 T; 0 U; 0 Other;
 XX
 Query Match 3.1%; Score 148.4; DB 4; Length 1023;
 Best Local Similarity 49.1%; Pred. No. 1.6e-14;
 Matches 483; Conservative 0; Mismatches 491; Indels 10; Gaps 3;
 XX
 QY 629 AAAATGAATAATGATTTTAAAGTGAAGATCATTTGGACAGCAATGCAATTCGACTG 688
 DB 4 AAACAAAAGTCTGCTTTTAAAGTCAAGATCTCTGGGACAGCTTTTAAAGTCAATTA 63
 QY 689 TCAGCAGATGATATATCACTTAATTGAGGACGATGATCAAGAAATTAACATATA 748
 DB 64 GCGGAAAATGACACAGAGTCTGATTTTGGGAAATTAACAGAGCAATGATGAAATA 123

QY 749 AACGTTACCGAAAAAATTTAAATTTTACCACATATCATCTACAGCAACATATAT 808
 DB 124 AATACAAAACACACCAACAAACATTTATTTACGAAATTTATCTTACTTCAATTCAG 183
 QY 809 GCAACGACATATATAGCAGAGATATATCTGACACAAATAC--ATGATATATCTTACTA 866
 DB 184 GCAACGACATATATAGCAGAGATATATCTGACACAAATAC--ATGATATATCTTACTA 243
 QY 867 TTCCATACACAAATATAGCAGATATATGACAAATATACAAACAAACGATATGTA 926
 DB 244 ACTAAAGCCATTCGTTAGTACAGAGAGTGGCCCAATTTAAACAAACCAAT 303
 QY 927 AAATATCTCAATATTTATTTAGTAAAGATATGAAATTTATCATCATCAATTTCCCA 986
 DB 304 ATATTTATATGCAATTAAGGCTCGAACAAGATCAATTAAGGATTTACAGAGTATA 363
 QY 987 GTGAATATGACAGAAATTTTACAAATATATCAATTTTATCTCTGTGTCAGATT 1046
 DB 364 GCGGAAAGATTCACAGAGA--AAACGCAAGGATGTGTTTATCTGTGTCAGAGC 421
 QY 1047 TTGCTAAGAAATGACAGAAATCTCTGATATATGATCTTGTGCTGATATTAAG 1106
 DB 422 ATGAGAGAGAGTGTGCTGTTATGATCAATTAACCAATTAACAGCAGAGAAATTTTG 481
 QY 1107 AACTGTGTAATCATTTGATGAAACAAATTAATGATGTTCTTAAATATATATACATC 1166
 DB 482 CAGACGCTGTATGATGCAAGAGTGTGTTATGATATATTAATTAATTAAGATG 541
 QY 1167 AAGATATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1226
 DB 542 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 601
 QY 1227 GAATTAATCGCTGAAAAAATTTTAAATATATGATGATGATGATGATGATGATGATG 1286
 DB 602 GAGGATNTCATGCTTAAAGCTTTGATGATGATGATGATGATGATGATGATGATG 661
 QY 1287 TGAATGAATTAATAACATATATATGCAAAAAATATCATATGATCTTCAATCTTA 1346
 DB 662 TACAGAAATTAATGCTGTTTAAAGGCTGCAATGG-----AGCAACCATTAACATTTA 715
 QY 1347 TTGTCATATGCTTCTGAGATCTATATATTAATGATGATGATGATGATGATGATGATG 1406
 DB 716 TTGCTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 775
 QY 1407 TGGCTTTGACATGAATATGAAAAAGTGAATTAATTAATTAATTAATTAATTAATTA 1466
 DB 776 GGGGTGGGATTAATTTATTTAGGAAAGGCAATTAATTAATTAATTAATTAATTAATTA 835
 QY 1467 TAAAGCTTTGAGGAAACCAATGATGATGATGATGATGATGATGATGATGATGATGATG 1526
 DB 836 GCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 895
 QY 1527 ATGTAAGTCAATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATG 1586
 DB 896 AAGTCGAAAGCAATCAACAGAAACATCTACATGCTTGTGTAATGACGAAGATGAC 955
 QY 1587 ATAAGCATATCAATATATAT 1610
 DB 956 AACAGCAGCAAAAGAAATCATCT 979
 XX
 RESULT 29
 AA551287
 ID AA551287 standard; DNA; 1026 BP.
 XX
 XX AA551287;
 DT 13-FEB-2002 (first entry)
 DE Enterococcus faecalis DNA for cellular proliferation protein #64.
 XX
 XX Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;

KW antibacterial; drug design.
 XX Enterococcus faecalis.
 XX
 PN NO200170955-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001WO-US009180.
 XX
 PR 21-MAR-2000; 2000US-0191078P.
 PR 23-MAY-2000; 2000US-0206848P.
 PR 26-MAY-2000; 2000US-0207727P.
 PR 23-OCT-2000; 2000US-0242578P.
 PR 27-NOV-2000; 2000US-0253625P.
 PR 22-DEC-2000; 2000US-0257931P.
 PR 16-FEB-2001; 2001US-0269308P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GU,
 PI Yamamoto RT, Xu HH;
 XX
 DR WPI: 2001-611495/70.
 DR P-PSDB; AAU33428.
 XX
 PT New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids.
 XX
 PS Claim 27; SEQ ID NO 3869; 511pp; English.
 XX
 CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the genes,
 CC their use in the discovery of novel antibiotics, the essential genes
 CC themselves and the discovery of proteins. The prokaryotes used are Escherichia
 CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
 CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
 CC useful for the identification of potential new targets for antibiotic
 CC development. The antisense nucleic acids can also be used to identify
 CC proteins used in proliferation, to express these proteins, and to obtain
 CC antibodies capable of binding to the expressed proteins. The proteins can
 CC be used to screen compounds in rational drug discovery programmes. The
 CC antisense nucleic acid sequence is also useful to screen for homologous
 CC nucleic acids which are required for cell proliferation in a wide variety
 CC of organisms. The present sequence encodes an essential prokaryotic
 CC cellular proliferation protein. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pat_sequences
 CC
 CC Sequence 1026 BP; 346 A; 172 C; 239 G; 269 T; 0 U; 0 Other;
 SQ
 Query Match 3.1%; Score 147.2; DB 4; Length 1026;
 Best Local Similarity 49.1%; Pred. No. 2,4e-14;
 Matches 480; Conservative 0; Mismatches 488; Indels 10; Gaps 3;
 QY 635 AAAATGAGTATTTAGTGCAGAGATCATTTGGACAGCAATATGCACTGTACACA 694
 DB 4 AAGTGCAGTGTATTAGTCCAGGTTCTGGGGAGACGTTTAGCTCAAGTATTACCGAA 63
 QY 695 CATGATATCATGTTAAGTATGGGACGCTGATCATAGAAATATACATATAACT 754
 DB 64 AATGGACAGAGGCTGTATTTGGGAAATTAACAGACCAATGATGAAATTAATCA 123
 QY 755 TACGAAAAAATTTAAATATTATCCACATATCACTACAGACAAATATATGCAAC 814
 DB 124 AAACACACCAACAAATATTATTCAGAAATTAATCTTACTTCAATTCAGCAAC 183
 QY 815 AGCAATATGACAGATATTATCTGCAACAATAC-ATGATTTATCTTAATCTTCC 872
 DB 184 ACTGATTAGCCACTGCTTAGTAGATGTGATGGGATCTGTTTGTGGTACCACTAAA 243
 QY 873 CACACATATTAAGCACAATATATGACACAATATCAACACAAACAGATATGTATAAATA 932

DB 244 GCCATTCGTTTACATGACACAGAAAGTGGCCCAACATTTAAAAACAAACCAATTATT 303
 QY 933 CTCCAAATTTAATTTAGTAAAGGATTCGAATATACATCACTCAAAATTTCCAGTAAA 992
 DB 304 CATGCAAGTAAAGGCTCGCAACAGGTACACATTAAGCGGATTTTCAGAAAGTATCGCGAA 363
 QY 993 TAGCAGAAAGAAATTTTCAATATATATCAATTTTATACCTCTGGTCCAAAGTTTCTA 1052
 DB 364 GAGATTCAGACGA-AAAACGGCAGAGATTTGGTTTATCTGTCACCAACCATGAC 421
 QY 1053 AAGAAATTCAGAAACATCTCTTGTAGTATAGTACTCTGTGTGATATTAAGAACTTG 1112
 DB 422 AAGAAAGTGTGCTCATGACATTCACCAATTCACAGACAGCAAGCAAAATTTAGCAACG 481
 QY 1113 GTGATCATTTGATGAACAATATAGTATGATGTTCTTAATAATATATACATCAAGATA 1172
 DB 482 CTGTGATGTGCAAGAGTTGTTATGATGATTAATTTAGAAATTTATACAAATGATATG 541
 QY 1173 TTATAGGTGACAGATTTGAGCTGCATTTAAAGAACATTAATTCATATGTAATAA 1232
 DB 542 TGATTTGGTGAAGAAAGGGGGCGGCTTTAAATAATATTATTTGGGGCGGACGCA 601
 QY 1233 TCGCTGAAAAAATTTAGTATTAATGCTGTGTACTGTATTAAAGCATGATG 1292
 DB 602 TTCAATGCTTAGGCTTTGGTATATATGCAAGCGCAATTAATGACTGTGTTAGCAG 661
 QY 1293 AAATTTAAACACTATATATATGCAAAAAATATCATATAGATCTTCATATCAATTAATGTC 1352
 DB 662 AATTTAGTCGTTTAGGGGTTGCCATGGG-----AGCAAAACCATTAATTTATGTT 715
 QY 1353 CATATGCTTTGAGATCTTATATTAATTAACATGATACACAGAAACATTACGCAATGCTT 1412
 DB 716 TAACTGTGTGGGATTTAATGTTATCTGTCACAGTCCATTCGCAAAATGGGGTG 775
 QY 1413 TTGCACATGAATAATGGAAGAGTGAATATTAATCATTAATAGATCACAACTTAAGC 1472
 DB 776 CGGTAATTTATTTAGAAAAAGCCATATAGTATGAAAGCTTATGAAAAATGGGCTGA 835
 QY 1473 TTGTGAAGAACCACTGCTGTAATAACCACTGATATCATTAAGCAAAAAAATTATAGTAG 1532
 DB 836 TTGTGAAGAGGTGTTCAACAACATTAAGCAACCTATGATTTGGAACAACAATCTGAAGTGC 895
 QY 1533 AACTACCAATTTGATATCTATTTACATTTATTTACATGAGATATATCTAGATTAAG 1592
 DB 896 AATGCGCAATTCAGAAACATTTCAATGATCTTTGTAATGAGAAAGATGTACACAG 955
 QY 1593 CCATATCAAAACATATTAT 1610
 DB 956 CAGCAAAAGAAATCATGT 973
 RESULT 30
 ABLJ32863/C
 ID ABLJ32863 standard; DNA; 6132 BP.
 XX
 AC ABLJ32863;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Human immune system associated gene SEQ ID NO: 836.
 XX
 KW Human; immune system disease; cytosine methylation; antiaesthetic;
 KW antiaesthetic; antiaesthetic; antiaesthetic; antiaesthetic;
 KW neuroprotection; anti-HIV; anticonvulsant; ophthalmological;
 KW antineoplastic; antineoplastic; antineoplastic; antineoplastic;
 KW antineoplastic; cancer; eye disease; arteriosclerosis; anaemia;
 KW acute myeloid leukemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
 KW ds.
 XX
 OS Homo sapiens.
 XX

PN WO200200928-A2.
 XX
 PD 03-JAN-2002.
 XX
 PF 02-JUL-2001; 2001WO-EP007537.
 XX
 PR 30-JUN-2000; 2000DE-01032529.
 XX
 PR 01-SEP-2000; 2000DE-01043826.
 XX
 PA (EPig-) EPIGENOMICS AG.
 XX
 PI Olek A, Piepenbrock C, Berlin K;
 DR WPI; 2002-130909/17.
 XX
 PT Nucleic acid comprising fragment of chemically modified gene, useful for
 PT diagnosis and treatment of diseases associated with abnormal cytosine
 PT methylation.
 PS
 PS Claim 1; SEQ ID NO 836; 32pp + Sequence Listing; German.
 XX
 CC The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention
 CC
 SQ Sequence 6132 BP; 1642 A; 28 C; 1246 G; 3216 T; 0 U; 0 Other;
 Query Match 3.1%; Score 147.2; DB 6; Length 6132;
 Best Local Similarity 44.2%; Pred. No. 2,6e-14;
 Matches 1069; Conservative 0; Mismatches 1318; Indels 31; Gaps 10;
 QY 281 AAGAAACAAGTATGATCAAGAAATACCTCTCTCTGTCATCCATATTTCCATTTAT 340
 Db 5037 AACAAACAACTTATCTTCAACACCTCTTAAATTTTCCCAAAATCCCTAAATAAT 4978
 QY 341 TTACAAATATCATAGAGTTCTTTATTCGATCATATTTAAATCAATCAAAACAAATA 400
 Db 4977 ATTAATAAATATTAATAA--TTATATAACACTTAATTAATTAATAATAATAATA 4921
 QY 401 TAGTACTATCTATTATATAAAAAATCCGATAGCTCTTAGCAAGAACTCAAAATATA 460
 Db 4920 AAAACATTAAGTCAACCTTAATAAAAAATTAATAAAAAATTAATACTATCTACAT 4861
 QY 461 TGAAGACCTTAATCTATTAATTAATTTCAAAATTTATTAATCTATCATATTTAGCAT 520
 Db 4860 TTATATACAC-AAACAATTAATTAATCTTAATTAATAACAAATTAATTAATAACAAA 4802
 QY 521 AATATCACTGATGACATGATTAACAATAGTACACATTTCTTAACACTAGATTAATATCA 580
 Db 4801 ACTACCAAAACAACTCAAAAAATTAATAAAAAATCCAAACAAATCTCACTCATCTA 4742
 QY 581 ATTAATATATTAATTTTGAAGTCTTCTGATTAATAATAGTAAATAATGAATAATC 640
 Db 4741 CTAATCAATATTAAC--TAAATTTCTTAACAAACAAATTAATAAAAAATTAATAAAA 4684
 QY 641 AGAATTTAGTGAAGATCATTTGGGACAGCAATAGCAATGCACTGACGACATGGT 700
 Db 4683 AAAATATCAATTAATTAATAAAAAATTAATACTATTTCTATTTCAAAATTAATACAT 4624
 QY 701 ATATCAATTAATTTAGGGAGCTGATCATAGAAATATTAACATATTAACACTTAACGA 760
 Db 4623 ATACCAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4564
 QY 761 AAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 820
 Db 4563 CAAATTTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4504
 QY 821 ATAGACGAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 880

Db 4503 AAAAAAATACTAAA-----AAACAATTCCTTTATTAATTAATCCCAAAAAATTAATAATAC 4448
 QY 881 TTACGACCAATATGATACAAAATTAACAACAAACAGCATATGTTAAATTAATCTCAATA 940
 Db 4447 ATAAACCTAATTTAAACCAAAAAATTAATAAAAAATTAATTAATTAATTAATTAATTA 4388
 QY 941 TTAATTTGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 1000
 Db 4387 CTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4328
 QY 1001 GAAATTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1060
 Db 4327 CAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4278
 QY 1061 GCAGAAATCTCTCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1120
 Db 4277 TCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4218
 QY 1121 TTGATGAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1180
 Db 4217 AAATCCAAAAAACAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4158
 QY 1181 GTACAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 1239
 Db 4157 TCTCAATTTCAAACTTAACAACTTAACAACTTAACAACTTAACAACTTAACAACTTA 4098
 QY 1240 AAAAAATTAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1299
 Db 4097 AAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4038
 QY 1300 AACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1359
 Db 4037 TTTAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3978
 QY 1360 TCTTGAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1419
 Db 3977 TCTTCAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3918
 QY 1420 AGAATTAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1477
 Db 3917 CTACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3858
 QY 1478 GAAAGAACAGTACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1537
 Db 3857 AAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3798
 QY 1538 CCAATTTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1597
 Db 3797 AATTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3738
 QY 1598 TCAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1657
 Db 3737 ACATTAATTAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3678
 QY 1658 GTCAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1717
 Db 3677 TACAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3618
 QY 1718 TAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1777
 Db 3617 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3558
 QY 1778 CTAGCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1837
 Db 3557 CAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3498
 QY 1838 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1897
 Db 3497 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3438
 QY 1898 CACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1957

Db 64774 ---ATCCACAACTTTTACGGGCTAAGTGTGATGACTTATCGTTACTTGTACAA 64718
 QY 1389 CAGAACATTCACGCAATATGCGTTTGGACTAGAAATAGAAAAAGGTGAATATAATA 1448
 Db 64717 GTGTGATTCACGTATATGGCGTGTGAAAACATCTCGGTAAAGCCAGAAATTTAGATG 64658
 QY 1449 CATTAAATAGATCAACAACCTTAAAGCTTGTGAAGAACAGTACTGTAATTAACCACTGATAT 1508
 Db 64657 AAGTTTAAAGAAAAATGGGTATGCTGTGAAGGTGTGACAGCCAAAGCAGTTGATG 64598
 QY 1509 CATTAGCAAAAAAATCTTAATAGTAACTACCAATTTGATCTATTATTAATTAATAC 1568
 Db 64597 GTTGGCGGAAAAAATAGATATTTGATATGCCAATTAACCAATTCGATTTTAT 64538
 QY 1569 ATGAGATATTA 1579
 Db 64537 TCGAGAAATMAA 64527

RESULT 32

AB067194
 ID AB067194 standard; DNA, 319630 BP.

AC AB067194;

DT 29-AUG-2002 (first entry)

DE Listeria innocua contig DNA sequence #7.

KW Antibacterial; Listeria; food contamination; mutational analysis;

KW infection; ds.

OS Listeria innocua.

PN WO200228891-A2.

PD 11-APR-2002.

PF 04-OCT-2001; 2001WO-FR003061.

PR 04-OCT-2000; 2000FR-00012697.

PA (INSP) INST PASTEUR.

PA (CNRS) CNRS CENT NAT RECH SCI.

PI Kunst F, Glaeser P;

DR WPI: 2002-332479/37.

XX New genomic sequences from Listeria species, useful for detection,

PT treatment and prevention of infection, also related polypeptides,

PT antibodies and modulators.

PS Claim 5; SEQ ID NO 7; 180bp; French.

XX The present invention relates to nucleic acid sequences (AB067188-

CC AB071212) from Listeria sp. The sequences are useful as probes and

CC primers for identification and/or detection of Listeria (e.g. as

CC contaminants in foods, or mutational analysis) and for analysis of gene

CC expression. Proteins encoded by the nucleic acid sequences can be used to

CC screen for compounds that modulate gene expression, replication and

CC pathogenicity of Listeria (potential therapeutic agents), also for

CC treating infections by Listeria, and are useful as immunogens in anti-

CC Listeria vaccines. Note: The sequence data for this patent did not form

CC part of the printed specification, but was obtained in electronic format

CC directly from WIPO at ftp.wipo.int/pub/published_pcf_sequences

XX Sequence 319630 BP; 105207 A; 55428 C; 66726 G; 92263 T; 0 U; 6 Other;

XX

Query Match 3.1%; Score 146.2; DB 6; Length 319630;
 Best Local Similarity 49.8%; Pred. No. 4.5e-14;
 Matches 484; Conservative 0; Mismatches 473; Indels 14; Gaps 4;

QY 613 GATATTAATAGTAAATAATGAAATCAGTATTTTATGCTGAGATCATTTGGCAGAC 672
 Db 51941 GAAAAATATATATACACAGAAAAAAGTGTCTATTTCTGGCGGTGAAAGTGGGAAACAG 52000
 QY 673 AATAGCAATTTGCACTGTGACACATGTGTATATACATTAATCTATGGGACGATCATG 732
 Db 52001 ACTGGACCTTGTCTGTGCTGATTAATATATTAACCAATTAATTTGGGAACTTAGATA 52060
 QY 733 AATATTAACATATATTAACATTAACCGAAAAATTTAAATATTTAACCAATATCATCT 792
 Db 52061 AATTTGAATGAATTAATGAATTCGACAGCAATATGCTATTTGGCAGATTAATTTT 52120
 QY 793 ACCAGCAACATATATATGCAACCAATATTAACGAAATTTATCTGACACAAATATCATG 852
 Db 52121 ACCAACTGAGTAAAGCAACATTAATCACTTATATAGTAAATGATGCTGAAT--- 52177
 QY 853 TATTATCTTAATCTATCTTACACAAATTTAGCAGCAATATGTACAAATATCAACACAA 912
 Db 52178 TGTGTGATTTGCTATTTCCAAATATGCAATGTGCTGTGTTGTAGCAGCTAAATGAAGC 52237
 QY 913 ACAGCAATATGTATTAATATCTCAATATTAATTTGTATGAAGTATGCAATATCATC 972
 Db 52238 GCTGAA-AGAACCACTATTTATGTCATGTTAGTAAAGGATTTGAACCAAGAAACAAATC 52296
 QY 973 ACTCAATTTCCAG---TGAATAGCAGAAATTTTACATATTAATCAATTTTAA 1028
 Db 52297 TTGAAATGTCAAGATTAATGAAGAAATGATGCTACAAAGTAAAGCTTTGTTG 52356
 QY 1029 TACTCTGTGTCGAATTTTGTCTAAAGAAATTCAGAAATCTCTCTGTATATATGATC 1088
 Db 52357 TTCTTCAGGGGCTGATGCTGAGTAAAGTATGCTCTGCTATCCAAAGCCTTTG 52416
 QY 1089 TTGCTGTATATTAAGAACTGTGATCATTTGATGAACAAATTAATATATGATGCTC 1148
 Db 52417 CTAGCTGTAAGTTTGAAGACAGCAATATGTCAGATGCTTTATTAATTAATTT 52476
 QY 1149 TAAATTAATATATCAATCAATCAATATTAATGATGATGATGATGATGATGATGATG 1208
 Db 52477 TACGTATTAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 52536
 QY 1209 TAATGCAATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1268
 Db 52537 TTAATGCAATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 52596
 QY 1269 CTGTTATTAATTAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1328
 Db 52597 CACTAATGACAGCGGAATGCTGAATTAATCTGCTGGAATGCTGCTGCTGCTGCTGCTG 52653
 QY 1329 TAGATCTTATATATTAATTTGTCATCATGCTTTGGAGATCTATATTAATATGATGATG 1388
 Db 52654 ---ATCCACAACTTTTACGGGCTTACGATGATGATGATGATGATGATGATGATGATG 52710
 QY 1389 CAGAACATTCACGCAATATGCGTTTGGACTAGAAATAGAAAAAGTGAATATTAATA 1448
 Db 52711 GTGTGATTCACGTATATGGCGTGTGAAAACATCTCGGTAAAGCCAGAAATTTAGATG 52770
 QY 1449 CATTAAATAGATCAACAACCTTAAAGCTTGTGAAGAAACAGTACTGTAATTAACCACTG 1508
 Db 52771 AAGTTTAAAGAAAAATGGGTATGCTGTGAAGGTGTGACACCAAGAGAGTTCATG 52830
 QY 1509 CATTAGCAAAAAAATCTTAATAGTAACTACCAATTTGATCTATTATTAATTAATTAAC 1568
 Db 52831 GTTGGCGGAAAAAATAGATATTTGATATGCCAATTAACCAATTCGATTTTAT 52890
 QY 1569 ATGAGATATTA 1579
 Db 52891 TCGAGAAATMAA 52901

RESULT 33

ABK28332/C
 ID ABK28332 standard; DNA, 11745 BP.

XX

AC ABK28332;
XX
XX DT 23-APR-2002 (first entry)
XX
DE DNA transcription associated complementary genomic DNA #103.
XX
XX
XX DNA transcription associated gene; peptide nucleic acid; PNA-oligomer;
KW PNA; cytosine methylation state; SNP; retroviral infection; gene; ds;
KW single nucleotide polymorphism; adenosine deaminase deficiency; cancer;
KW viral infection; Sezary syndrome; haematological disorder; tuberculosis;
KW immunological disorder; Werner syndrome; developmental disorder;
KW psoriasis; Rieger's syndrome; neurologic disorder; erythropoiesis;
KW neurodegenerative disorder; Waardenburg syndrome; Niemann-Pick disease;
KW myelodysplastic syndrome; myocardial infarction; hyperextension; arthritis
KW angiogenesis; congenital heart disease; HDR syndrome; gene therapy;
KW polyglutamine disorder; solid tumour.
XX
XX
XX Unidentified.
XX
XX
XX WO200192565-A2.
XX
XX
XX 06-DEC-2001.
XX
XX
XX 06-APR-2001; 2001WO-EP003973.
XX
XX
XX 06-APR-2000; 2000DE-01019058.
XX 07-APR-2000; 2000DE-01019173.
XX 30-JUN-2000; 2000DE-01032529.
XX 01-SEP-2000; 2000DE-01043826.
XX
XX
XX (EPIC-) EPIGENOMICS AG.
XX
XX
XX Olek A, Piepenbrock C, Berlin K;
XX
XX
XX MPI; 2002-090046/12.
XX
XX
XX New nucleic acids or oligomers, useful for diagnosing or treating
PT diseases associated with DNA transcription, e.g. immunological disorders,
PT ptner syndrome, psoriasis, myocardial infarction, solid tumors or
PT cancer.
XX
XX
XX Claim 1; SEQ ID NO 206; 32pp; English.
XX
XX
XX The invention relates to a nucleic acid, which comprises a segment of chem
CC chemically pre-treated DNA of genes associated with DNA transcription from
CC one of 346 sequences, and an oligomer, in particular an oligonucleotide
CC or peptide nucleic acid (PNA)-oligomer that hybridises to or is identical
CC to the chemically pre-treated DNA of genes associated with DNA
CC transcription. The set of oligomer probes are useful for detecting the
CC cytosine methylation state and/or single nucleotide polymorphisms (SNPs)
CC in a chemically pre-treated genomic DNA. The nucleic acids are useful for
CC diagnosing or treating diseases associated with DNA transcription
CC (particularly with the methylation status), e.g. adenosine deaminase
CC deficiency, viral infection, retroviral infection, Sezary syndrome,
CC haematological disorders, immunological disorders, Werner syndrome,
CC tuberculosis, developmental disorders, psoriasis, Rieger's syndrome,
CC neurological disorders, neurodegenerative disorders, Waardenburg
CC syndrome, Niemann-Pick disease, myelodysplastic syndrome, myocardial
CC infarction, hyperextension, angiogenesis, erythropoiesis, congenital heart
CC disease, HDR syndrome, arthritis, polyglutamine disorders, solid tumours
CC or cancer. Sequences ABK28127-ABK28472 represent DNA transcription
CC associated genomic DNA molecules of the invention. Note: The sequence
CC data for this patent did not form part of the printed specification but
CC was obtained in electronic format directly from the European Patent
CC Office
XX
XX
XX Sequence 11745 BP; 3183 A; 270 C; 2383 G; 5908 T; 0 U; 1 Other;
SQ
Query Match 3.1%; Score 145.4; DB 6; Length 11745;
Best Local Similarity 43.5%; Pred. No. 5,2e-14;
Matches 968; Conservative 0; Mismatches 1241; Indels 18; Gaps 6

| | | | | | |
|---|---|--|------|---------------------------------------------------------------------|------|
| D | b | | 4950 | -----TTTAAATATAAATTATACACAAGAAATTAATCAAAATTCAAAATACACTTAAGA | 4897 |
| O | y | | 1466 | CTAAGCGTTGTGAAGAACGAGTAGCTGTAAAAACCATGATATCATTTGCAAAAAAAGCTT | 1525 |
| D | b | | 4896 | ATTAACAATAAAGCTTACTTAATATATATATATATATATAAATATAAAAAATAAATATGCA | 4837 |
| O | y | | 1566 | AATGAGAGACTACCAATTGGCATATCTATTTTGCAATTATATACATGAGAAATATATCCTA | 1585 |
| D | b | | 4836 | AATTTACATCTA-AACTTTTAAGCTAAAAAACCAGATTAATATATATCTTATTTCTTAA | 4778 |
| O | y | | 1586 | GATAAAGCCATATCCAACATATTAATCTTAGTCTATCATCCTTTGATAATTTTCAACACAT | 1645 |
| D | b | | 4777 | TAAAAAAATTTATATATAAAAAAAAACTAAATTTTAATTAATAAATAATCAATAATATAT | 4718 |
| O | y | | 1646 | GATTAATTAAGCGTCAAAAAATATGAAATACACCGTTGCACCATATATTAATAAGACAGCA | 1705 |
| D | b | | 4717 | ATCATTAATCTGATTTTTHACCGGTATCAAAATATCTATTAACGATCTTAAAAATATTA | 4658 |
| O | y | | 1706 | AATACCCATCGTTAAACATCCACTAAGATGTACATATTTACATPACAACTTAAGATATA | 1765 |
| D | b | | 4657 | AATPACCACTATATATATTAATCTTAA--AACTTAATAAAAAAATAAATAAATAAAGCATT | 4600 |
| O | y | | 1766 | ACTTAAATTTCCGTGTAATTAATTAATAAGAAATCCGATCCGATCTCAATTAACCA | 1825 |
| D | b | | 4599 | TTTTAAATATACATCACATPAAAAAATATATTTTAAACCTTAATAATTAATCTGAAATCACCTA | 4540 |
| O | y | | 1826 | TTGTTATCAGATTATATAATAAGTCAAATTAACTTTTACTATTTATTAATAATATGTT | 1885 |
| D | b | | 4539 | AAAAAAAAAAAAACAATAAAAAAAAAAAAAAAAAAAC-----CAAAAACAAAAACTAAATCAC | 4485 |
| O | y | | 1886 | TGCACACTTTTTCACATCATGATGATATATGAACATPATTTTACATPAATAATACATAT | 1945 |
| D | b | | 4484 | TCTACATTTAAATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT | 4425 |
| O | y | | 1946 | AAAAAGAACCAATACATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT | 2005 |
| D | b | | 4424 | TAAATATAAAAAAAAAAACTPACAAATATATATCAAAATPACATTAATAAAAAAAAAAACTTTC | 4365 |
| O | y | | 2006 | AGATGCTCGTCATPATTAATTTATCATCTGTCGACATATTTCAAACATTAATAATAGTATA | 2065 |
| D | b | | 4364 | AAAACAATAAATAATATCAATTAATCATATATATCTTTTAAATCTTAAACAAACTTAA | 4305 |
| O | y | | 2066 | AACTGTTTTTTTATGCCATTAGTTCTTAATTAATAATGCTTAATTAACCAAAAAGTTTAA | 2125 |
| D | b | | 4304 | ATCATTTATATTTTATCAATATATTAATCAATTAATTAATTTATTAATTAATTAATAAATCTTTTAA | 4245 |
| O | y | | 2126 | ATACACATCTCTGCACAAATATGAAGATACATATACAGTCAATTTAAATTAACCAAT | 2185 |
| D | b | | 4244 | TTAAATTAATAAATAAATAATATATCCATTAATTAATTAATTAATTAATTAATTAATTAATTAAT | 4185 |
| O | y | | 2186 | TTCTPATATTAACCATCTACTCTTAACATAGAACCAATAGT-ATATCATCTTAAGTAA | 2244 |
| D | b | | 4184 | AAAAACAACAATATTAACAACAATTTAATAAAAAATTTCACTAAATTAATAAAAAAACA | 4125 |
| O | y | | 2245 | ATATCCCTACAGCATGTATCAATTAATTAATAAATAAATAATGCTATGTATATTAATCAATTA | 2304 |
| D | b | | 4124 | TAAAGGTATATTAATAATTAATAATTAATAAATAAATAAATTTCTTAAATTAATTAATTAATA | 4065 |
| O | y | | 2305 | CNAATTTCCAACTTCAACAAACAAACCTCAAAAGTTCTTATTAACAAATTCAGAAATTAAT | 2364 |
| D | b | | 4064 | CAACTCTATATATCAATTAATAAATAATTAACCAATTAATAAAAAAACAATAATTAACAAACA | 4005 |
| O | y | | 2365 | ATATATCTTACCAACACATATGTTTATCAAAATCTATTAATACATCATATCCTTTTAAANAAG | 2424 |
| D | b | | 4004 | AAAAAAATATCTTCAAAAAATTAATAATCTTTTAAAAAAACAATTAATAAATAATTAATAAC | 3945 |
| O | y | | 2425 | ATGACCTTCACATTCATCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT | 2484 |
| D | b | | 3944 | ACAAATTAATAAATACTAACCTTAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAA | 3885 |
| O | y | | 2485 | CAATTCAGAAATTTTCAATTTTACGGGTATAGTGAAGCTTCCATTCACAAATTCAGC | 2544 |
| D | b | | 3884 | AAAAATATATTAATTAATTTAAAAACAATAATTAATTAATTAATTAATTAATTAATTAATTAAT | 3825 |

[illegible]


```

Db      7268 AAAACATCTAAATAATAATAATAATAATAAACCCTTTTAAATTAATTCMAATAA 7209
Qy      2747 ATAACTATTAGTATTTGCTTTAAGTTACTGGCAATTCATGACTGCAATGCTATTAA 2806
Db      7208 AAATAAATTAAACAACTTAATAAACAATCTTCAAAAAAATAATAATAATAA 7149
Qy      2807 GCTTCATCATCTAATATACATGAGTACAAACCTATTCTTTTACAGATGCAATTCATATA 2866
Db      7148 AAACAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 7093
Qy      2867 TACCTACACAACTTCGGTATATCTGACAAATATCTACTAAGACGGTACTCTAATGGA 2926
Db      7092 CTCCTTATTTTCCAAAAATAAATACTAAATATCTCTATAAAAAAATAATAATAA 7033
Qy      2927 AGGACATTTAATCTAATATAATTAATCCTCAAAAACCTTCTTACTCTCACTTCA 2986
Db      7032 ATAAAAATAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 6973
Qy      2987 ATATCTTTGAGAGAGACACAAATATCTCACAATCTACTAACAAGAACTTACTATT 3046
Db      6972 AACCTTACCTAACCACTAACAAATATCTAATAATATATAATAATAATAATAATACTATCT 6913
Qy      3047 TCCTCTAT 3053
Db      6912 TTCAAAAT 6906

```

RESULT 36
AB067070/c
ID AB067070 standard; DNA: 5611 BP.

```

XX      28-AUG-2002 (first entry)
XX      Human angiogenesis associated polynucleotide SEQ ID NO 100.
XX      Human: angiogenesis; methylation; eye disease; glaucoma; tumor;
XX      inflammation; rheumatoid arthritis; diabetic retinopathy; antidiabetic;
XX      muscular degeneration; inflammatory bowel disease; Crohn's disease;
XX      antineutrotic; antidiabetic; antidiabetic; antipsoriatic;
XX      antidiabetic; antidiabetic; ds.
XX      Homo sapiens.
XX      W0200246454-A2.
XX      13-JUN-2002.
XX      06-DEC-2001; 2001MO-EP014320.
XX      06-DEC-2000; 2000DE-01061338.
XX      (EPIC-) EPIGENOMICS AG.
XX      Schacht O;
XX      WPI; 2002-500450/53.
XX      New nucleic acid fragments from chemically treated angiogenesis-
XX      associated genes, useful for determining methylation status, e.g. in
XX      diagnosis or treatment of cancer.
XX      Claim 1; SEQ ID NO 100; 41pp + Sequence Listing; German.
XX      The invention relates to a nucleic acid (I) comprising a segment of 18
XX      bases of chemically pretreated DNA of angiogenesis-associated genes (II)
XX      having sequences (AB066971-AB067178) or their complements. (I), also
XX      related oligomers, are used to evaluate the methylation status and/or
XX      single-nucleotide polymorphisms, in angiogenesis-related genes, for
XX      diagnosis and treatment of eye diseases, proliferative retinopathy,
XX      neovascular glaucoma, solid tumors, inflammation, rheumatoid arthritis,

```

CC diabetic retinopathy, macular degeneration caused by neovascularisation,
CC psoriasis, arteriosclerosis, inflammatory bowel diseases, ulcers and
CC Crohn's disease. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 5611 BP; 1825 A; 38 C; 926 G; 2822 T; 0 U; 0 Other;

Query Match 3.0%; Score 141; DB 6; Length 5611;
Best Local Similarity 44.8%; Pred. No. 2.3e-13;
Matches 820; Conservative 0; Mismatches 980; Indels 31; Gaps 6;

```

Qy      700 TATATCAGTTACTTATGAGGACGTGATCATAGAAATTTACATATATTAACACTTACCG 759
Db      2866 TATTTTACTACTTTTAAAAATAATAAAAAAGATTTATCTAATTAATAAATTTCC 2887
Qy      760 AAAAAATTTAAATTTTATCCCATATTCATCTCCAGACACATATATGACACAGCA 819
Db      2806 TCTAATAATAAAAAAATAAACAACATCTATTATTATTTCCATCAATTAATCAA 2747
Qy      820 TATAGACGAGATTTATCTGACCAACATATGATATCTTACTATCTCTACACACA 879*
Db      2746 AATTAATAATCTAATATATCTTCAAACTTTTATATCTTTAATTAATTAATAAATCT 2687
Qy      880 ATTACGACCATATGTAACAAATATACACACAAACAGCATATGTGTAATAATCTCCAT 939
Db      2686 ATTAAAAAATAAAAAATTTTACATTCATTAATTTACTATTA-ATATTAATAAATACATAT 2628
Qy      940 ATTAATTTGTGTAAGATTCGAAATTTACATCTCAAAATTTCCAGTGAATAGAGA 999
Db      2627 TTTAAATTTAATAAAAAAATAATATATCTTAAAAATTTTATTTCAAAAATAATA 2568
Qy      1000 AGAAATTTTACAAATATATCCAAATTTTATCTCTGTCGCAAGTTTGTGTAAGAAT 1059
Db      2567 CAACCTCTTACCAACAATATATATTTTATTTAATTTTAAAAAATAAAAAATAA 2508
Qy      1060 TGAGAACATCTCTCTGTGTATGATAGTCTGCGTGATATTAAGACTGGTGAATC 1119
Db      2507 AAAAAAATAATATATTAATTTTAAATCTCAATTTCAATTAATAAATTTTCTAAAA 2448
Qy      1120 ATTGATGAACAATATATGATATGTTCTTAAATTAATATACA----- 1164
Db      2447 ATATACCAATTAATAATAAATAAATAATTAATCTTAAATTTCTAAAAATAAATAATA 2388
Qy      1165 ---TCAAGATTTATAGCTGTACAGATTGAGCTGCCTTAAAGACATATGCAATTC 1221
Db      2387 ATCAAAACCTATCTTAATATTAATAAACAATAATATATTAATATATCTTACTTA 2328
Qy      1222 ATGTGAATAATCGCTGGAATAAATTTAGGTAATATATGCTGTGCTATGTAATAA 1281
Db      2327 ATATTAATTTTAAAAAATAAATAATCTAATTAATAAATTTTATCTTTAACACCTA 2268
Qy      1282 AGGATGAATGAATTTAAACACTATATATAGCAAAAAATCATTCATATAGATCTTCAAC 1341
Db      2267 ACATTAATTAACAACATTAATTTCAATCAACCTTAACAACTTAATTTATTAACCTTAA 2208
Qy      1342 ATTATTTGTCATCATCTATGCTTTGAGATCTAATATTAATTAACATGTAACAACAGATTGACG 1401
Db      2207 ACAAAATCTCCCTTTAATCTTTAATCTTTTATTAATAATA---AAAAAATAATTA 2151
Qy      1402 CAATATGAGCTTTTGGATAGAAATAGAAAAAGTGAATAATATATATATATAGTCA 1461
Db      2150 TCTATCCCACTTATTTTCAATTAATCTAATAAATAATCAATTAATAATTAATTAAGCAATCA 2091
Qy      1462 CAACCTTAAGCTTTGTGAAGAACCACTACTGTATTAACCACTGATATCATTAAGCAAAAA 1521
Db      2090 CTTAAAAAATCTAATAATCTTACATCATATTAATAATAATAAATAAATAATATATCTA 2031
Qy      1522 ACTTAATGTAGAACTACCAATTTGCAATATCTATTTTACATTTATTAACATAGAAATATATC 1581
Db      2030 TATTAATACCCCAACATTAACATCTCCATCAACCTACTACTACCAATCAATCAATATAT 1971
Qy      1582 ACTGATTAAGCACTATCAACAATATATCTTATGCTATATCATCTTTGATTAATTTGACA 1641

```

Db 1970 TTAATACCACTATACATATCCAAATATATTTTCCAAACATCTTAAATTAATTAATTT 1911
 QY 1642 CAATGATTAATTAACGGTCGAAAAATATGAAATPACACGTTCCACATATATTAATTAAGCA 1701
 Db 1910 CAATATCTTAATTAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1851
 QY 1702 GTCAAAT-ACCAGTCGTAAACATCCAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTA 1760
 Db 1850 TAAAAATTAACATATCTTACATCTCCAAATTTCTAATTTCTAATTTCTAATTTCTAATTT 1791
 QY 1761 GATTAACCTTAATTTCTAGCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1820
 Db 1790 ACAAAATCAAAACATATCATCTTAATCCAACTTCTAATTTCTAATTTCTAATTTCTAATTA 1731
 QY 1821 ACCAATGTTATAGTTATTAATTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTA 1880
 Db 1730 ATTATATTTATATTAATTAATTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTA 1671
 QY 1881 ATGTTGCACTTTTCTAATCATCATATATTAATTAAGTAAGTAAGTAAGTAAGTAAGTAAGTA 1940
 Db 1670 AATATTAACCTTTTATATTAATTAATTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTA 1611
 QY 1941 CTAAATTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTA 2000
 Db 1610 ATACATATTAATTAATTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAAT 1552
 QY 2001 TACATAGATGCTGCAATATTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTA 2060
 Db 1551 ATATTAATTAACATATCTTATTAATTAATTAATTTCTAATTTCTAATTTCTAATTTCTAATTT 1492
 QY 2061 TGAATTAAGTTTATTTATGCTTATTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAAT 2120
 Db 1491 AATTAATTAATTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTA 1432
 QY 2121 TTAATTAACATTTCTGCAATATTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTT 2180
 Db 1431 TAAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTT 1372
 QY 2181 CAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTT 2233
 Db 1371 AATTAATTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTT 1312
 QY 2234 ACCTAATTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTT 2293
 Db 1311 ACTTAATTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTT 1252
 QY 2294 TTAATTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTT 2353
 Db 1251 AATTAATTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTT 1192
 QY 2354 TCAGATTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTT 2413
 Db 1191 AATTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTA 1132
 QY 2414 TTTTAAATTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTT 2473
 Db 1131 ACTTAATTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTT 1072
 QY 2474 GAAATTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTT 2504
 Db 1071 ATTAATTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTTCTAATTT 1041

RESULT 37
 ID ADC91772
 AC ADC91772 standard; DNA; 1062 BP.
 XX
 XX
 XX
 XX
 DE E. faecium DNA sequence SEQ ID 1399.

XX
 KW ds; gene; urinary tract infection; bacteraemia; endocarditis; wound;
 KM abdominal-pelvic infection.
 XX
 OS Enterococcus faecium.
 XX
 PN US6583275-B1.
 XX
 PD 24-JUN-2003.
 XX
 PF 30-JUN-1998; 98US-00107532.
 XX
 PR 02-JUL-1997; 97US-0051571P.
 PR 14-MAY-1998; 98US-0085598P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Doucette-Stamm LA, Bush D;
 XX
 DR WPI: 2003-799836/75.
 DR P-PSDB; ADC95426.
 XX
 PT New isolated nucleic acid derived from Enterococcus faecium encoding an
 PT Enterococcus faecium polypeptide useful for detection, prevention and
 PT treatment of a pathological condition resulting from a bacterial
 PT infection.
 XX
 PS Example 1; SEQ ID NO 1399; 243bp; English.
 XX
 CC The invention relates to an isolated nucleic acid derived from
 CC Enterococcus faecium encoding an Enterococcus faecium polypeptide having
 CC one of 10 fully defined sequences given in the (or comprising 40
 CC sequential nucleotides chosen from any of the nucleic acids, its
 CC complement or sequences hybridizing to it). Also included are a
 CC recombinant vector comprising the nucleic acid operably linked to
 CC transcription regulatory element, a cell comprising the vector and a
 CC single-stranded probe comprising the nucleic acid. The nucleic acids are
 CC chosen from 3654 disclosed sequences encoding 3654 disclosed proteins.
 CC The nucleic acids is useful for diagnosing pathological conditions
 CC resulting from E. faecium bacterial infection (e.g. urinary tract
 CC infection), bacteraemia, endocarditis, wounds and abdominal-pelvic
 CC infection) and for screening drugs such as agonists and antagonists. The
 CC nucleic acid is useful for recombinant production of Candida albicans -
 CC derived peptides or antisense polypeptides. Pharmaceutical compositions
 CC and vaccines containing the nucleic acid are useful for preventing or
 CC treating Enterococcus faecium infections. The present sequence represents
 CC one if the disclosed E. faecium nucleic acids.
 XX
 SQ Sequence 1062 BP; 362 A; 204 C; 246 G; 250 T; 0 U; 0 Other;
 XX
 Query Match 2.9%; Score 139.8; DB 9; Length 1062;
 Best Local Similarity 49.6%; Pred. No. 3.6e-13;
 Matches 484; Conservative 0; Mismatches 467; Indels 24; Gaps 4;
 QY 645 TTTTAAAGTCAAGATCAATTTGCGACGAAATGCAATTTGCACTGTCGACGATGATAT 704
 Db 56 TTTTAAAGTCAAGATCAATTTGCGACGAAATGCAATTTGCACTGTCGACGATGATAT 115
 QY 705 CAGTTAACTATGAGGAGCGATCATAGAAATATTAACATTTAAACATTTCCGAAAAA 764
 Db 116 AGCTTGGATTTGGGACATCTGCTTCACAGTAGAAGAAATCAACACCAACACACA 175
 QY 765 ATTTAAATTTATTTACCGACATATCTTACGACACAACTATATGCAACGACATATAG 824
 Db 176 ATCAAGCTTATTTACCGACATATCTTACCGACATATCTTACCGACATATCTTACCGACATATG 235
 QY 825 ACAGATATATGACCAACATATATATTTATCTTAACTATTTCTTACCAACATATAC 884
 Db 236 AAGATGAGTAAAGCGACATGCGTCTATCTCTATCTTCAAGAAAGCAATATC 292
 QY 885 GACCATATGTACCAAAATGCAACACAAACAGCATATGTTAAATTAATCTCCATATTTAA 944
 Db 293 GTCTGTGCAAAAGAACTAGTCCCTAAA-----TTGATTAACAAAGCCAGTATCATC 346

QY 1209 TAATTCATTCATGATGATATATCGCTGGAAAAATTTAGTAAATATGCTGTGCTA 1268
 Db 1260 TATTCACATAGGTGACAGAAATTTCTGACGGCCTTGTTATGCGGTAATGCTAAAGCG 1319
 QY 1269 CTGTATACTAAAGCATGATGAATTAATAAACAATATATATAGCAAAAAATCATTCOA 1328
 Db 1320 CACTAATGACTCGCGGATGGAGAAATCATCTGCTGCTGGTTC-----TGTGGCT 1373
 QY 1329 TATATCTTCAATATTAATTTGCTTCATCATGCTTGGAGATCTAATATTAATGACAA 1388
 Db 1374 CTATCCGCAAACTATTTAGATGACAGGATTTGGTACCTTAATCTGTTACTGTACTA 1433
 QY 1389 CAGAACATTCAGCAATATGCTTTGACCTAGAAATAGAAAGGTAGAAATATATAA 1448
 Db 1434 GTGTGATTTACCTTAATTTGCGCGTCTGCTGTAATATGCTAGTAAAGCGAAATTTAGATG 1493
 QY 1449 CATTAAATAGTCAAACTTAAGCTTTGTAAGAAACAGTACTGTAACCACTGATAT 1508
 Db 1494 AAGTATTAGAAAAAATGGGATGCTGCTGTAAGGTTGAAACAGCTAAAGCAGTACATG 1553
 QY 1509 CATTAGCAAAAAAATCTTAATGTAAGACTACCAATTTGCAATCTATCTAATTTATATAC 1568
 Db 1554 GTTGGCGGAAAAAATAGATATGATATGATGCCAATTTACGAAATGATTTAGCGGATTTAT 1613
 QY 1569 ATGAGAAATATA 1579
 Db 1614 TCGAAAAATATA 1624

RESULT 39

ABL3307/C
 ID ABL3307 standard; DNA; 6175 BP.

AC ABL3307;

DT 26-MAR-2002 (first entry)

DE Human immune system associated gene SEQ ID NO: 1280.

XX Human; immune system disease; cytosine methylation; antiasthmatic;
 KM antiarteriosclerotic; antihaemic; cytosine; noctropic;
 KM neuroproliferative; anti-HIV; anticonvulsant; ophthalmological;
 KM antihemorrhagic; antidiabetic; antidiabetic; antiproliferative;
 KM antineoplastic; cancer; eye disease; arteriosclerosis; anaemia;
 KM acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KM neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
 XX ds.

OS Homo sapiens.

PN W0200200928-A2.

PD 03-JAN-2002.

PF 02-JUL-2001; 2001WO-EP007537.

PR 30-JUN-2000; 2000DE-01032529.

PR 01-SEP-2000; 2000DE-01043826.

PA (EPIC-) EPICENOMICS AG.

PI Olek A, Piepenbrock C, Berlin K;

DR WPI; 2002-130909/17.

XX Nucleic acid comprising fragment of chemically modified gene, useful for
 PT diagnosis and treatment of diseases associated with abnormal cytosine
 PT methylation.

PS Claim 1; SEQ ID NO 1280; 32bp + Sequence Listing; German.

CC The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences

CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/vulserative bowel
 CC diseases. The present sequence is a gene of the invention
 XX

SO Sequence 6175 BP; 1506 A; 154 C; 1357 G; 3158 T; 0 U; 0 Other;

Query Match 2.9%; Score 139; DB 6; Length 6175;

Best Local Similarity 45.8%; Pred. No. 5.2e-13;

Matches 800; Conservative 0; Mismatches 910; Indels 35; Gaps 8;

QY 740 ACACATATAACCTTACCGAAAAATTTAAATTTTACCCCATATGATCTACAGAC 799
 Db 6002 ATAAATATATAAACAATAATATATATAAATTTAAATATACAAATATAATCTTAATAT 5943
 QY 800 AACATATATGCAACAGCAATATAGACAGATATTTCTGACCAACATACATGATATATC 859
 Db 5942 TATATATAAACAACGATATAACATATAATATATATACATAAACAATAATATATATACAT 5883
 QY 860 TTAATCTTTCTTACACACCAATTTACGACCATTTGTACCAATATACACCAACAGCAT 919
 Db 5882 ACGTAAACAAATATATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5823
 QY 920 ATGTGTAATAATCTCCATATATTTTGTAGTAAGATGTAAGTATGCAATATACATCACTCAA 979
 Db 5822 TAAATATATAAATAAATCTAATATTAACGATATATAATATAATATAAACA-----ACTATAAAT 5767
 QY 980 TTTCGCCAGTAATAGCAGAGAAATTTTACAATATATATCAATTTTATCTCTCTGCT 1039
 Db 5766 ATAAACATATATATAAACAATAATATATATATATATAAACAATATATAAACAATATATA 5707
 QY 1040 CCAAGTTTCTTAAAGAAATTTGAGAACATCTCTCTGTATGATGATGATGATGATGAT 1099
 Db 5706 TACAAATATATATAAACAATAATATATATATAATATAATATAATATAATATAATATA 5647
 QY 1100 AATTAAGAACTGCTGATCATTTGATGAAGAAACAATATATATATATATATATATAT 1159
 Db 5646 TATTAATATATAAACAATATATATATATATAAACAATATATAATATAATATAATATAT 5587
 QY 1160 TACATCAAGATATTTAGGTGACAGATTTGAGCGCATTAAGAACAATATATGCAAT 1219
 Db 5586 TACATATATAAATAATTTAACAATATATATAAACAATATATAAACAATATATAA 5527
 QY 1220 GCATGTGAATATTCGCTGAAAAAATTTAGTAATATATGCTGCTCTATATTAAT 1279
 Db 5526 AAATTAATATATAAATAATATAAATAAATAATATATAAATAAACAATATATATAA 5467
 QY 1280 AAAGCATGAATGAATTTAAACATATATATATGCAAAAAATGATCAATATAGATCTTCAT 1339
 Db 5466 AAAGCTAAACATATATATAAATAATATATATATGCTAAATATAATATAATATAATATAT 5411
 QY 1340 ACAATTAATGCTCATCATGCTCTTGGAGATCTAATATTAACATGTAACAACAGAACTCA 1399
 Db 5410 ACACAAATATATAAACAACAACATATATATATATATATATATATATATATATATAT 5351
 QY 1400 CGCAATGCTCTTTGACAGTAAGAAATAGGAAAGGTAGAAATTAATCATATATAGAT 1459
 Db 5350 AAATATATATGATATAAACAATATAAAGTAATATATATATAAACAATATATATATAT 5291
 QY 1460 CAGAACCTTAAGCTTGTGTAAGAACAGTACTGTAAACAACACTGATATCATTTAGCAAAA 1519
 Db 5290 ATATATATATAAATAAATAAATAAATATATATATATATATATATATATATATATAT 5231
 QY 1520 AAATTAATGATAGACATCAATTTGCAATATATATATATATATATATATATATATATAT 1579
 Db 5230 ATATATATAAACAATATATATATAAACAACAATATATATATATATATATATATATAT 5171
 QY 1580 TCACTGATTAAGCATATCAACATATATATCTTATGATATCACTCTTGTATATTTCAA 1639
 Db 5170 TATATAAATAAATAAACAATATATATATATAAACAATATATATATATATATATATAT 5112

| | | | |
|--------------------------------------------------------------------------|------|-----------------------------------------------------------------|-------|
| QY | 1640 | CACCAATGTAATAAATACGGTCAAAAATATAGAAATCAAGT-TCACGACATATATATAAA | 1698 |
| Db | 5111 | AAATATAACCAATATATATATATAAACCTAATATATAATTTATAACAAATATATATAACA | 50323 |
| QY | 1698 | GCAGTCAAAATACCCATCGTTACATCCACTAAGTAGTATGACATATTACATACAACTT | 17588 |
| Db | 5051 | TAAATATATAAACCTTAATATATAAAATATATATAACAAATATATAATAAACATATATAATTA | 49922 |
| QY | 1759 | AAGTATAACTTAAATTTCTGCTAATATATATAACAAATTCCTATCTCATCTCAAT | 18188 |
| Db | 4991 | AAAAATACATTAATTAATATACCGTATATATAAACAACTAAATTAATCTAAATATATATCA | 49322 |
| QY | 1819 | TACCAATCTGTATACGATTTATATAAGTCAAAATTAACCTTTACTTATTTATATAATA | 18788 |
| Db | 4931 | TATATATATAATTAATACATTTAAATATAATATAATATACATATAATTAATAT-...AAATA | 48755 |
| QY | 1879 | ATATGTTTGACACTTTTTCATCATCATGATATATATATGAACAATATTTTACATATAAAT | 19388 |
| Db | 4874 | AAATATAAACATTAATTAATTAATATAACCAATATATATATAATTAATTAATATATAACA | 48155 |
| QY | 1939 | AACATATAAAGAACCAATACATATATTAATTTCTAAACACCTATTTTCTACGTAT | 19988 |
| Db | 4814 | AATACATATTAATTTTAAACATATATAAATAACCAACATTAATTAATTAATTCGTATA | 47555 |
| QY | 1999 | ATTACATGATGTCCTGCATTAATAATTTCTATCATCTCTGCATATTTTCAACATATAAA | 20588 |
| Db | 4754 | AATACCAATATATATAAACAATTTATTAATTAATTAATTAATTAATTAATTAATTAAT | 47600 |
| QY | 2059 | TATGATTAACGTGTTTTTTTATGCCATTAAGTCTTAATATAAATTGCTAATTAACCAAAA | 21188 |
| Db | 4699 | ATAAACATATATACATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT | 46400 |
| QY | 2119 | GGTTAAAAATACACTTCCTGCACATATGAAGATACATACATACAGCTCAAAATATAATTA | 21788 |
| Db | 4639 | TATAATATATACAAAT-...AAATATATTAATTAATTAATTAATTAATTAATTAATTAAT | 45897 |
| QY | 2179 | AACAAATTTCTAATTAATTAACACCTTCTTAACAATGAACCAATAGTATTCATACCTA | 22388 |
| Db | 4586 | CACAAACATATATATTAATTAACATATTAATTAATTAATTAATTAATTAATTAATTAAT | 4537 |
| QY | 2239 | AGTAGATATCCCTACACAGATGTATCAATATATAAGAAAAATCGTTATGTATTTATA | 22988 |
| Db | 4536 | ATATATACATATATATTAACAACATATATATATATCAATTAAGCTATTAACGTATA | 4477 |
| QY | 2299 | CAATTACAAATTTCCACTTCAACACAGAAATCTCAAGTCTTATTAACAAATTTCAAG | 23588 |
| Db | 4476 | TAAATATTAACATTAATTAATTAATTTACATTAATTTTAAACATAATTAATTAATTAATTA | 4417 |
| QY | 2359 | AATATATATATCTTCACCAACACATAGTTATCAAAATCTATATATATACATCATATCTTTT | 24188 |
| Db | 4416 | AAAAATACGTATATTAACACGTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT | 4357 |
| QY | 2419 | AAAAAGATGACTTCACATTCATCATTAATTAATTAATTAATTAATTAATTAATTAATTAAT | 24788 |
| Db | 4356 | AAAAAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA | 4297 |
| QY | 2479 | ATTAC 2483 | |
| Db | 4296 | AATAC 4292 | |
| RESULT 40 | | | |
| ABQ66998/c | | | |
| ID ABQ66998 standard; DNA; 37515 BP. | | | |
| XX ABQ66998; | | | |
| XX 28-AUG-2002 (first entry) | | | |
| XX Human angiogenesis associated polynucleotide SEQ ID NO 28. | | | |
| XX Human; angiogenesis; methylation; eye disease; glaucoma; tumour; | | | |
| XX inflammation; rheumatoid arthritis; diabetic retinopathy; anti-leads; | | | |

QY 584 AATATATATATATTTTGAAGTCTTCTGATATATATAGTAAAAATGAAAAATCAGT 643
Db 18430 AAAATATATATATATATATATATATATATATATATATATATATATATATATATAT 18371
QY 644 ATTT-----TAGGCGAGATCATTTGGCAGAGATAGCAATTGACGCTGCAGCAAG 699
Db 18370 ATTTATATATATATATATATATATATATATATATATATATATATATATATATAT 18311
QY 700 TATATGATTAATCTTATGAGGAGCTGATCATAGAAATATATACATATATACATTAACG 759
Db 18310 AAAATCTAAAAACAATAATTT-CTTATCTCATTAATCTTATATATATATATATATAT 18252
QY 760 AAAATATATATATATATATATATATATATATATATATATATATATATATATATAT 819
Db 18251 CAAAATATATATATATATATATATATATATATATATATATATATATATATATAT 18192
QY 820 TATATGAGATATATATCTGACACATATATATATATATATATATATATATATATAT 879
Db 18191 TATATATATATATATATATATATATATATATATATATATATATATATATATAT 18132
QY 880 ATTATGACCATATATATATATATATATATATATATATATATATATATATATATAT 939
Db 18131 TAAAAATATATATATATATATATATATATATATATATATATATATATATATAT 18072
QY 940 ATTAATTTGATATATATATATATATATATATATATATATATATATATATATATAT 999
Db 18071 ATTAATCTTATATATATATATATATATATATATATATATATATATATATATAT 18012
QY 1000 AGAATTTTATATATATATATATATATATATATATATATATATATATATATAT 1059
Db 18011 AAAATTTTATATATATATATATATATATATATATATATATATATATATATAT 17952
QY 1060 TGAGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1119
Db 17951 ACCCAATATATATATATATATATATATATATATATATATATATATATATAT 17869
QY 1120 ATTATATATATATATATATATATATATATATATATATATATATATATATATAT 1179
Db 17869 AATATATATATATATATATATATATATATATATATATATATATATATATATAT 17839
QY 1180 TGTATGATATGAGTATATATATATATATATATATATATATATATATATATATAT 1239
Db 17838 TCTTATATATATATATATATATATATATATATATATATATATATATATATAT 17779
QY 1240 AAAATTTTATATATATATATATATATATATATATATATATATATATATATAT 1299
Db 17778 AAAATTTTATATATATATATATATATATATATATATATATATATATATATAT 17720
QY 1300 AACATATATATATATATATATATATATATATATATATATATATATATATATAT 1359
Db 17719 TCAAAATATATATATATATATATATATATATATATATATATATATATATATAT 17660
QY 1360 TCTTATATATATATATATATATATATATATATATATATATATATATATATAT 1419
Db 17659 TCTTATATATATATATATATATATATATATATATATATATATATATATATAT 17600
QY 1420 AGAATATATATATATATATATATATATATATATATATATATATATATATATAT 1479
Db 17599 TCAAAATATATATATATATATATATATATATATATATATATATATATATATAT 17540
QY 1480 AGAATATATATATATATATATATATATATATATATATATATATATATATATAT 1539
Db 17539 TCAAAATATATATATATATATATATATATATATATATATATATATATATATAT 17481
QY 1540 AATTTGATATATATATATATATATATATATATATATATATATATATATATATAT 1599
Db 17480 AATTTGATATATATATATATATATATATATATATATATATATATATATATATAT 17421
QY 1600 AATATATATATATATATATATATATATATATATATATATATATATATATATAT 1659
Db 17420 TCAAAATATATATATATATATATATATATATATATATATATATATATATATAT 17361
QY 1660 CAAAATATATATATATATATATATATATATATATATATATATATATATATATAT 1719

Db 17360 ATTCAATATATATATATATATATATATATATATATATATATATATATATATATAT 17301
QY 1720 ACATATATATATATATATATATATATATATATATATATATATATATATATATAT 1779
Db 17300 AACATATATATATATATATATATATATATATATATATATATATATATATATAT 17241
QY 1780 AGTATATATATATATATATATATATATATATATATATATATATATATATATAT 1839
Db 17240 ATCAATATATATATATATATATATATATATATATATATATATATATATATATAT 17183
QY 1840 TAAATATATATATATATATATATATATATATATATATATATATATATATATAT 1899
Db 17182 TTAATATATATATATATATATATATATATATATATATATATATATATATATAT 17123
QY 1900 CTATATATATATATATATATATATATATATATATATATATATATATATATATAT 1959
Db 17122 CTCAATATATATATATATATATATATATATATATATATATATATATATATATAT 17063
QY 1960 CATATATATATATATATATATATATATATATATATATATATATATATATATAT 2019
Db 17062 ACTATATATATATATATATATATATATATATATATATATATATATATATATAT 17003
QY 2020 AATATATATATATATATATATATATATATATATATATATATATATATATATAT 2079
Db 17002 AATATATATATATATATATATATATATATATATATATATATATATATATATAT 16943
QY 2080 GCCATATATATATATATATATATATATATATATATATATATATATATATATAT 2139
Db 16942 ATCTTATATATATATATATATATATATATATATATATATATATATATATATAT 16886
QY 2140 AACATATATATATATATATATATATATATATATATATATATATATATATATAT 2199
Db 16885 ACCTATATATATATATATATATATATATATATATATATATATATATATATATAT 16826
QY 2200 CTTATATATATATATATATATATATATATATATATATATATATATATATATAT 2259
Db 16825 TAAATATATATATATATATATATATATATATATATATATATATATATATATAT 16766
QY 2260 GTATATATATATATATATATATATATATATATATATATATATATATATATATAT 2319
Db 16765 TTAATATATATATATATATATATATATATATATATATATATATATATATATAT 16709
QY 2320 CAAATATATATATATATATATATATATATATATATATATATATATATATATAT 2379
Db 16708 AAAATATATATATATATATATATATATATATATATATATATATATATATATAT 16649
QY 2380 ACATATATATATATATATATATATATATATATATATATATATATATATATATAT 2439
Db 16648 TAAATATATATATATATATATATATATATATATATATATATATATATATATAT 16589
QY 2440 ATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2499
Db 16588 GTCTATATATATATATATATATATATATATATATATATATATATATATATATAT 16529
QY 2500 CAATATATATATATATATATATATATATATATATATATATATATATATATATAT 2559
Db 16528 C---TGGCATATATATATATATATATATATATATATATATATATATATATATAT 16473
QY 2560 ACCTATATATATATATATATATATATATATATATATATATATATATATATATAT 2619
Db 16472 ACTATATATATATATATATATATATATATATATATATATATATATATATATATAT 16413
QY 2620 AGTATATATATATATATATATATATATATATATATATATATATATATATATATAT 2679
Db 16412 AACATATATATATATATATATATATATATATATATATATATATATATATATATAT 16353
QY 2680 TATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2739
Db 16352 TTTATATATATATATATATATATATATATATATATATATATATATATATATATAT 16293
QY 2740 CATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2798

[illegible]

| | | | |
|----|-------|--------------------------------------------------------------------|-------|
| OY | 3857 | TTGGACATTTTCAAAATCCGTGGCATAATTAATCTCAAAATCAATGAATGAATCTGGCAACA | 3916 |
| Db | 15155 | TACTTCCTTTAAATTTTTCATATC-----TAAAAAGTATTTATTAATTCGGTAT | 15106 |
| OY | 3917 | CAATGTATCCCTTTAAAGTACACTCATTAATCTACACACACATTAAGATCACACCC | 3976 |
| Db | 15105 | CCAAATACCTCTCATPACTCATTAATAAATTTTCTACTACACACTCGTACCCCAAAATTC | 15046 |
| OY | 3977 | AATATATTAATTCATTATTTTTTTACTTAATAATTTTAAACATATTAATAAAATTAAGTAC | 4036 |
| Db | 15045 | CTAATAAATCTATCCATTAATTAATAAAATTAATAAATATTTAAATAAATTAATTAATTAACAA | 14986 |
| OY | 4037 | TATATTAAGTAATTTTCAAGCATAGTGTCCATTAATTCACAAAAGAAAAATTAATAACA | 4096 |
| Db | 14885 | ATPAACCTTAACAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAACA | 14926 |
| OY | 4097 | CTTAATTAATAAACTTATCATCAACTTTGTAAACCTATAGTAACTAATTAATTAATAATTA | 4156 |
| Db | 14925 | AAATTAATAAAACATTAATTAATTAATAAACTTAACCTTTAAATAAAATTAATAATAAATA | 14866 |
| OY | 4157 | TTTCAATATTTTACCTCCCTGGAAACAGCATGTCTATAGTTTCCATACATAAAATTC | 4216 |
| Db | 14865 | TTTCAAAAATAAAACT-----AAAAAAATCTCTAAATTAACCATTAATAAAACAAAAA | 14813 |
| OY | 4217 | ACTTGAAGCTTTAATAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAG | 4276 |
| Db | 14812 | AAATATATTTTCAAAAACATTAATTAACATTTCCCAAAAAATTAATTAATTAATTAATTAATAA | 14753 |
| OY | 4277 | TAAATTTATATTTTAAACCATAGACAAAATTAATTCCTCCACATGACTTCAACATATGT | 4336 |
| Db | 14752 | CTAACCTTTATTAATAAAAAAAAACAAATTAATTAACCATTTTAAAAAATAATAAAAAAATAA | 14693 |
| OY | 4337 | GAATAGTGAACCATACACCAACAAGAGATTAATTCATTGACTTACTTCCACGTATAAATC | 4396 |
| Db | 14692 | AAATTAATAATTTTAAATAACCATTAATACTATTAATTCATTAACACAAATTAACAAATTAATAAT | 14633 |
| OY | 4397 | TAGATGATTAACGACCATCTCCATGTATTTGAAGTCAAAATCATCCGATATTTACAAT | 4456 |
| Db | 14632 | TTCTAAATTTACACATCTTCACTTAATTAATTAATTTTTCAAATTTCTACCGATTAATTAATA | 14573 |
| OY | 4457 | ATCAACAATTAATGTCCAAAACATCAATCAATGATAGAAATTTT | 4498 |
| Db | 14572 | CTAAACGAATTAATTAATAAATTAATAAATTAATAAATTAAT | 14531 |

Search completed: April 5, 2004, 13:26:52
Job time : 1773 secs

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2004, 10:55:36 ; Search time 17898 Seconds

(without alignments) 11502.928 Million cell updates/sec

Title: US-10-081-051-52

Sequence: 1 gatccctgctactactatgag.....ccacattgactaatgac 4750

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapept 1.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

1: gb_da:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_ov:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_hlg_hum:*
31: em_hlg_inv:*
32: em_hlg_other:*
33: em_hlg_mus:*
34: em_hlg_pla:*
35: em_hlg_rtd:*
36: em_hlg_mam:*
37: em_hlg_vtc:*
38: em_sy:*
39: em_hgo_hum:*
40: em_hgo_mus:*
41: em_hgo_other:*

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|------------|---------------------|
| 1 | 4750 | 100.0 | 4750 | AF308670 | AF308670 Cowdria r |
| 2 | 4750 | 100.0 | 4750 | AX685245 | AX685245 Sequence |
| 3 | 1416 | 29.8 | 1416 | AX685249 | AX685249 Sequence |
| 4 | 984 | 20.7 | 984 | AX685247 | AX685247 Sequence |
| 5 | 597 | 12.6 | 597 | AX685250 | AX685250 Sequence |
| 6 | 371 | 7.8 | 371 | AX685246 | AX685246 Sequence |
| 7 | 357.6 | 7.5 | 279110 | REPX03 | REPX03 |
| 8 | 338.6 | 7.1 | 12477 | AE008640 | AE008640 Rickettsia |
| 9 | 231 | 4.9 | 23906 | AF176314 | AF176314 Zymomonas |
| 10 | 219.6 | 4.6 | 14211 | AE011408 | AE011408 Leptospira |
| 11 | 216 | 4.5 | 11400 | AB080665 | AB080665 Wolbachia |
| 12 | 207.8 | 4.4 | 34980 | AX344555 | AX344555 Sequence |
| 13 | 204 | 4.0 | 349174 | AB63522 | AB63522 Mligeswo |
| 14 | 188.6 | 4.0 | 4362 | PVGLNABC | PVGLNABC |
| 15 | 186.4 | 3.9 | 8056 | AX599046 | AX599046 Sequence |
| 16 | 184 | 3.9 | 4601 | DMU11584 | DMU11584 Drosophila |
| 17 | 184 | 3.9 | 13517 | DMU37541 | DMU37541 Drosophila |
| 18 | 183 | 3.9 | 172816 | AC093899 | AC093899 Homo sapi |
| 19 | 182.2 | 3.8 | 14867 | AE001398 | AE001398 Plasmodi |
| 20 | 182 | 3.8 | 30205 | AE015944 | AE015944 Clostridi |
| 21 | 181.4 | 3.8 | 145992 | AC134919 | AC134919 Homo sapi |
| 22 | 180.2 | 3.8 | 12168 | AC127383 | AC127383 Homo sapi |
| 23 | 175.8 | 3.7 | 12277 | AE010598 | AE010598 Fusobacte |
| 24 | 174.4 | 3.7 | 333750 | AP004827 | AP004827 Staphyloc |
| 25 | 173.8 | 3.7 | 104992 | AC005504 | AC005504 Plasmodi |
| 26 | 173.8 | 3.7 | 169546 | AC004157 | AC004157 Plasmodi |
| 27 | 173.8 | 3.7 | 250421 | AE014849 | AE014849 Plasmodi |
| 28 | 173.4 | 3.7 | 301278 | AE015939 | AE015939 Clostridi |
| 29 | 172.8 | 3.6 | 301550 | AP003134 | AP003134 Staphyloc |
| 30 | 172.8 | 3.6 | 303414 | AE015938 | AE015938 Clostridi |
| 31 | 172.8 | 3.6 | 346900 | AP003362 | AP003362 Staphyloc |
| 32 | 170.8 | 3.6 | 282183 | CJ11168X4 | CJ11168X4 Sequence |
| 33 | 170.6 | 3.6 | 1425 | AR377574 | AR377574 Sequence |
| 34 | 169.2 | 3.6 | 10861 | AE007680 | AE007680 Clostridi |
| 35 | 169 | 3.6 | 3900 | AY309266 | AY309266 Borrelia |
| 36 | 168.8 | 3.6 | 13877 | AC120883 | AC120883 Homo sapi |
| 37 | 168.8 | 3.6 | 349751 | PFMAL4P3 | PFMAL4P3 Plasmodi |
| 38 | 168.6 | 3.5 | 996 | AX622666 | AX622666 Sequence |
| 39 | 168.4 | 3.5 | 67970 | PFMAL1P3 | PFMAL1P3 Plasmodi |
| 40 | 167 | 3.5 | 19517 | DMU37541 | DMU37541 Drosophila |
| 41 | 166.8 | 3.5 | 175559 | AC146101 | AC146101 Homo sapi |
| 42 | 166.4 | 3.5 | 13124 | AE008105 | AE008105 Agrobacte |
| 43 | 166.4 | 3.5 | 13147 | AE008070 | AE008070 Agrobacte |
| 44 | 166 | 3.5 | 254050 | PFPA929358 | PFPA929358 Plasmodi |
| 45 | 165.8 | 3.5 | 10277 | AE001176 | AE001176 Borrelia |

ALIGNMENTS

RESULT 1
LOCUS AF308670
DEFINITION AF308670 4750 bp DNA linear BCT 03-OCT-2001
Cowdria ruminantium clone 19hw hypothetical integrase/recombinase
gene, partial cds; hypothetical glycerol 3-phosphate dehydrogenase
and hypothetical nitrogen assimilation regulatory protein genes,
complete cds; and hypothetical cell division protein ftsQ gene,
partial cds.
ACCESSION AF308670.1 GI:15811157
VERSION AF308670
KEYWORDS
SOURCE
ORGANISM
Bacterial; Proteobacteria; Alphaproteobacteria; Rickettsiales;
Anaplasmataceae; Ehrlichia.

QY 961 CGAAATTCATCACTCAAAATTTCCAGTGAATAGCAGAAATTTTACATAATAATCC 1020
Db 961 CGAAATTCATCACTCAAAATTTCCAGTGAATAGCAGAAATTTTACATAATAATCC 1020
QY 1021 AATTTTATCTCTCTGTCGCAAGTTTTCGTAAGAAATTCAGAACATCTTCCCTGAG 1080
Db 1021 AATTTTATCTCTCTGTCGCAAGTTTTCGTAAGAAATTCAGAACATCTTCCCTGAG 1080
QY 1081 TTTAGTCTCTGTCGTCGTAATTAAGAACTTGTAATCATTTGATGAACAATAGTAA 1140
Db 1081 TTTAGTCTCTGTCGTCGTAATTAAGAACTTGTAATCATTTGATGAACAATAGTAA 1140
QY 1141 TGAATGTTCAAAATTAATATACATCAAGATTTATAGTGTACAGATTTGAGCTGAT 1200
Db 1141 TGAATGTTCAAAATTAATATACATCAAGATTTATAGTGTACAGATTTGAGCTGAT 1200
QY 1201 AAGAGACATTAATTCGCAATTCGATGGAATATGCTGGAAAAATTTAGGTAAATGC 1260
Db 1201 AAGAGACATTAATTCGCAATTCGATGGAATATGCTGGAAAAATTTAGGTAAATGC 1260
QY 1261 TGTGCTACTGTTATTAAGGATGATGAAATTTAAACATATATATAGCAAAAA 1320
Db 1261 TGTGCTACTGTTATTAAGGATGATGAAATTTAAACATATATATAGCAAAAA 1320
QY 1321 TCATTCATATAGATCTTCAATCAATTAATGTCATCATGTCCTGAGATCTAAATTTAC 1380
Db 1321 TCATTCATATAGATCTTCAATCAATTAATGTCATCATGTCCTGAGATCTAAATTTAC 1380
QY 1381 ATGTCAACAGACATTCACGCAATATGCTTTTGAATAGAAATAGGAAAAAGGTAGAAA 1440
Db 1381 ATGTCAACAGACATTCACGCAATATGCTTTTGAATAGAAATAGGAAAAAGGTAGAAA 1440
QY 1441 TATATAATCATTAATATAGATCAACACCTTAAGCTTGTGAGAAACAGATCTGTAACC 1500
Db 1441 TATATAATCATTAATATAGATCAACACCTTAAGCTTGTGAGAAACAGATCTGTAACC 1500
QY 1501 ACTGATATCATTAGCAAAAAAATTAATGTAAGAACTAACCAATTTGCATATCTTAA 1560
Db 1501 ACTGATATCATTAGCAAAAAAATTAATGTAAGAACTAACCAATTTGCATATCTTAA 1560
QY 1561 TTTATTAATGAGATATATATCAATGATTAAGCAATCAACATTAATTTAGTCTAT 1620
Db 1561 TTTATTAATGAGATATATATCAATGATTAAGCAATCAACATTAATTTAGTCTAT 1620
QY 1621 CATCTTTGATTAATTTCAACACATGATTAATTAACGTCAAAAATATAGATCAACGT 1680
Db 1621 CATCTTTGATTAATTTCAACACATGATTAATTAACGTCAAAAATATAGATCAACGT 1680
QY 1681 TCACACATATATAAAGAGAGCTCAAAATCCATGCTTTACATCACTAAATATATGTAAC 1740
Db 1681 TCACACATATATAAAGAGAGCTCAAAATCCATGCTTTACATCACTAAATATATGTAAC 1740
QY 1741 ATATTACATTAACAACTTAAGTATTAATTAATTTCTAGCTAATATATTAACAAAT 1800
Db 1741 ATATTACATTAACAACTTAAGTATTAATTAATTTCTAGCTAATATATTAACAAAT 1800
QY 1801 CCTATCTCTATCTCAATTAACCAATTTGATGATGATTAATTAATTAAGCAATTAAC 1860
Db 1801 CCTATCTCTATCTCAATTAACCAATTTGATGATGATTAATTAATTAAGCAATTAAC 1860
QY 1861 TTAATTAATTAATTAATTAATTTGTCGACACTTTTCACTACATGATATATATGAA 1920
Db 1861 TTAATTAATTAATTAATTAATTTGTCGACACTTTTCACTACATGATATATATGAA 1920
QY 1921 CATATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1980
Db 1921 CATATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1980
QY 1981 ACCTATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2040
Db 1981 ACCTATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2040
QY 2041 CATATTTTAAACATTAATTAATTAATTAATTAATTTTATATGCTTATGTTCTTAATAA 2100
Db 2041 CATATTTTAAACATTAATTAATTAATTAATTAATTTTATATGCTTATGTTCTTAATAA 2100
QY 2101 TTGCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2160
Db 2101 TTGCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2160
QY 2161 TACAGCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2220
Db 2161 TACAGCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2220
QY 2221 AATAGTATATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2280
Db 2221 AATAGTATATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2280
QY 2281 TCGTTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2340
Db 2281 TCGTTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2340
QY 2341 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2400
Db 2341 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2400
QY 2401 ATACATCATATCTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2460
Db 2401 ATACATCATATCTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2460
QY 2461 CAATGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2520
Db 2461 CAATGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2520
QY 2521 ACCTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2580
Db 2521 ACCTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2580
QY 2581 AGTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2640
Db 2581 AGTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2640
QY 2641 TTTAGCACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2700
Db 2641 TTTAGCACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2700
QY 2701 TGTGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2760
Db 2701 TGTGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2760
QY 2761 TGTGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2820
Db 2761 TGTGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2820
QY 2821 TACATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2880
Db 2821 TACATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2880
QY 2881 CGGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2940
Db 2881 CGGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2940
QY 2941 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3000
Db 2941 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3000
QY 3001 AGACACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3060
Db 3001 AGACACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3060
QY 3061 TCCCTTCGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3120
Db 3061 TCCCTTCGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3120
QY 3121 ATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3180
Db 3121 ATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3180

| | | | |
|----|------|---------------------------------------------------------------|------|
| Db | 3121 | TTCTATAAATAACGACCATGATTTGGTTCCTATATTCGAATATGAGAGGACTCT | 3185 |
| Oy | 3181 | ATGAGACATATATTTATTAATCTTTCCTCACTACCAAAATATTTAAACGAATATTAATAGC | 3240 |
| Db | 3181 | ATGAGACATATATTTATTAATCTTTCCTCACTACCAAAATATTTAAACGAATATTAATAGC | 3240 |
| Oy | 3241 | TGGTAGCATAGTAGTACATAGATATATAAGAGTATCATACCCTTGATTTTATG | 3300 |
| Db | 3241 | TGGTAGCATAGTAGTACATAGATATATAAGAGTATCATACCCTTGATTTTATG | 3300 |
| Oy | 3301 | TATTAGCCTAGCAATCTACTTCCCTTCCCAACACCTGGGGAACACGAATAGATATACACT | 3360 |
| Db | 3301 | TATTAGCCTAGCAATCTACTTCCCTTCCCAACACCTGGGGAACACGAATAGATATACACT | 3360 |
| Oy | 3361 | CGATGTAGTACCTGCTTTTATTAATCAATCTTCGAATTAAGTATTAACAGGGAGTTAC | 3420 |
| Db | 3361 | CGATGTAGTACCTGCTTTTATTAATCAATCTTCGAATTAAGTATTAACAGGGAGTTAC | 3420 |
| Oy | 3421 | GACATTTTCATTAATCCCAATAGCGATTTTAAACATCATCTTTCCTAGCTAATACAC | 3480 |
| Db | 3421 | GACATTTTCATTAATCCCAATAGCGATTTTAAACATCATCTTTCCTAGCTAATACAC | 3480 |
| Oy | 3481 | AGACTCTATAGCTCTCTTTAACACACTTAATCTTCCTGTGAAAAGGCTTTTCTAT | 3540 |
| Db | 3481 | AGACTCTATAGCTCTCTTTAACACACTTAATCTTCCTGTGAAAAGGCTTTTCTAT | 3540 |
| Oy | 3541 | ATAATCATTAAGAGCCATATGACAGACTTTTAAACGAGTAGGCAATATTAACATGCCACT | 3600 |
| Db | 3541 | ATAATCATTAAGAGCCATATGACAGACTTTTAAACGAGTAGGCAATATTAACATGCCACT | 3600 |
| Oy | 3601 | AATCATTAATAACAGGCAATAAGATACCTTTCTTTAAGCTTTTCCAGTACCTTAATCC | 3660 |
| Db | 3601 | AATCATTAATAACAGGCAATAAGATACCTTTCTTTAAGCTTTTCCAGTACCTTAATCC | 3660 |
| Oy | 3661 | ATCAATATCAGATCCTCTTTAAACATATATCCAAATATATCAACATCAGGCTTTTTCATA | 3720 |
| Db | 3661 | ATCAATATCAGATCCTCTTTAAACATATATCCAAATATATCAACATCAGGCTTTTTCATA | 3720 |
| Oy | 3721 | AGCCATCTTGATCGCGGATTAACCATCAACTGCTAATTTAGTACATATTAATCATCACT | 3780 |
| Db | 3721 | AGCCATCTTGATCGCGGATTAACCATCAACTGCTAATTTAGTACATATTAATCATCACT | 3780 |
| Oy | 3781 | TAAATATCTTTTATTAATATTCATATATCAACTCATCATCAACAACTATATACCTGGA | 3840 |
| Db | 3781 | TAAATATCTTTTATTAATATTCATATATTCATATATTCATATATTCATATATTCGAA | 3840 |
| Oy | 3841 | AATATATCAATCTTTCCTTGGACATTTCAAAATCCTGTGCATTAATTACCTACAATCTAA | 3900 |
| Db | 3841 | AATATATCAATCTTTCCTTGGACATTTCAAAATCCTGTGCATTAATTACCTACAATCTAA | 3900 |
| Oy | 3901 | TGAATATCTGCAAAACACATGTAATCCTTTAAAGAACCTCATATATCTACACACACCA | 3960 |
| Db | 3901 | TGAATATCTGCAAAACACATGTAATCCTTTAAAGAACCTCATATATCTACACACACCA | 3960 |
| Oy | 3961 | CTAATGATCAACACCAATATATTAATTCATATTTTCTATTAATTTAAATTAACAATAT | 4020 |
| Db | 3961 | CTAATGATCAACACCAATATATTAATTCATATTTTCTATTAATTTAAATTTAAACAATAT | 4020 |
| Oy | 4021 | TATATAAATTAAGTATCTATATTAATCTAATTTATACAGATAGGTTTCTATATATCCAA | 4080 |
| Db | 4021 | TATATAAATTAAGTATCTATATTAATCTAATTTATACAGATAGGTTTCTATATATCCAA | 4080 |
| Oy | 4081 | AAGAAAAATTAATAACCTTAATTAATAAACTTATCATCAACTTTGTAACCTATATGTAAT | 4140 |
| Db | 4081 | AAGAAAAATTAATAACCTTAATTAATAAACTTATCATCAACTTTGTAACCTATATGTAAT | 4140 |
| Oy | 4141 | AATTAATTAATAATTAATTCATATATTAATTAATTAATTAATTAATTAATTAATTAATTT | 4200 |
| Db | 4141 | AATTAATTAATAATTAATTCATATATTAATTAATTAATTAATTAATTAATTAATTAATTT | 4200 |
| Oy | 4201 | TCCATATCTTAATAACCTCACTGGAAGTTTATAGATTTCAATATTAACAGGCAACAGAT | 4260 |
| Db | 4201 | TCCATATCTTAATAACCTCACTGGAAGTTTATAGATTTCAATATTAACAGGCAACAGAT | 4260 |

| | | | |
|----|------|----------------------------------------------------------------|------|
| QY | 4261 | AAGAACTAATATTTGGTAATTTTATATTTATTTAAACGACATGACAAATTAATATCCACAAT | 4330 |
| Db | 4261 | AAGAACTAATATTTGGTAATTTTATATTTAAACGACATGACAAATTAATATCCACAAT | 4330 |
| QY | 4321 | GACTATCAACATATGTATATGATGAACCATPACCACCAACAGATTTATCATTAAGTA | 4380 |
| Db | 4321 | GACTATCAACATATGTATATGATGAACCATPACCACCAACAGATTTATCATTAAGTA | 4380 |
| QY | 4381 | CTTACCGTATATTAATCTAGATAGATTACCCACCATCTCATGTATGAAGTCAAAATCAT | 4440 |
| Db | 4381 | CTTACCGTATATTAATCTAGATAGATTACCCACCATCTCATGTATGAAGTCAAAATCAT | 4440 |
| QY | 4441 | CCCGTATATTTACAAATTTATCAACATTAATATGTCCAAACATATCAATATGAAATTTTGT | 4500 |
| Db | 4441 | CCCGTATATTTACAAATTTATCAACATTAATATGTCCAAACATATCAATATGAAATTTTGT | 4500 |
| QY | 4501 | TATCATGATACCAATTTTGGCAAAAGCAGATATATCTTGAACGTATCTGTAAGAATTTTG | 4560 |
| Db | 4501 | TATCATGATACCAATTTTGGCAAAAGCAGATATATCTTGAACGTATCTGTAAGAATTTTG | 4560 |
| QY | 4561 | GTAACAGCTCTTACTAGAGCATTTTGAATCCATGAGATGACTATATTTTGTTC | 4620 |
| Db | 4561 | GTAACAGCTCTTACTAGAGCATTTTGAATCCATGAGATGACTATATTTTGTTC | 4620 |
| QY | 4621 | TCAATATAGCCAAAGGACAAAAAAGATAGACCTGATCAACAGATTTACGTATTCAT | 4680 |
| Db | 4621 | TCAATATAGCCAAAGGACAAAAAAGATAGACCTGATCAACAGATTTACGTATTCAT | 4680 |
| QY | 4681 | CAGAACTTACATATTCATTAACCATCAATATTTTATCAACAGTAATTCACAAATTGA | 4740 |
| Db | 4681 | CAGAACTTACATATTCATTAACCATCAATATTTTATCAACAGTAATTCACAAATTGA | 4740 |
| QY | 4741 | CTAATGTATC 4750 | |
| Db | 4741 | CTAATGTATC 4750 | |

| | | | |
|----|----|--------------------------------------------------------------|-----|
| QY | 1 | GATCCTGCTAATACCTTAGGTAATGTTTTTTTAAAGATGCGCCTAGATAATAAATAAACT | 60 |
| DB | 1 | GATCCTGCTAATACCTTAGGTAATGTTTTTTTAAAGATGCGCCTAGATAATAAATAAACT | 60 |
| QY | 61 | GCTTGATGCTAATATAGTACATCGGATAGATAAGAAAAAATTTTCCAAGCAGAC | 120 |

Db 61 GCTGACGTCTCTATATAGTACATACGTGATGATTAACGAAAAATTTCTCAAGACAC 120
Qy 121 AGTACCTTGATGATGTAAGTCTTCACTCACTGATGATCTAGAGCAACATTTT 180
Db 121 AGTACCTTGATGATGTAAGTCTTCACTCACTGATGATCTAGAGCAACATTTT 180
Qy 181 CTCAATCTGCAATTTTACATGACAAATATTTATGTAACAGAACCTCTGACTC 240
Db 181 CTCAATCTGCAATTTTACATGACAAATATTTATGTAACAGAACCTCTGACTC 240
Qy 241 TTATGCAAGAACTCTATTAATCTTCAAGGCTCTTACATAGAAACATGATTCAA 300
Db 241 TTATGCAAGAACTCTATTAATCTTCAAGGCTCTTACATAGAAACATGATTCAA 300
Qy 301 GAATACCTCTTCTGTTCCATCCATATTTTCCATATTTTCAATATCATAGAGTCT 360
Db 301 GAATACCTCTTCTGTTCCATCCATATTTTCCATATTTTCAATATCATAGAGTCT 360
Qy 361 TTATTCGACATCATATTTAAATGATCAACAAATATATAGTATCTATTAATA 420
Db 361 TTATTCGACATCATATTTAAATGATCAACAAATATATAGTATCTATTAATA 420
Qy 421 AAAAATCCTCAATAGTCTTTAGCAGAACTACAAATATTAATGACACTTATTTAA 480
Db 421 AAAAATCCTCAATAGTCTTTAGCAGAACTACAAATATTAATGACACTTATTTAA 480
Qy 481 TATATTTCAATTTATTTATCTACTCATATTTTACATATATCATCTGACATGAT 540
Db 481 TATATTTCAATTTATTTATCTACTCATATTTTACATATATCATCTGACATGAT 540
Qy 541 TAAACATAGTAAACATTTCTTAACTAGATATATATACATATATATATATATTTT 600
Db 541 TAAACATAGTAAACATTTCTTAACTAGATATATATACATATATATATATATTTT 600
Qy 601 GACTGCTTTCTGATATTAATTTAGTAAATGAAAAATGAAATATGATTTTAGTGCAGAGATC 660
Db 601 GACTGCTTTCTGATATTAATTTAGTAAATGAAAAATGAAATATGATTTTAGTGCAGAGATC 660
Qy 661 ATTGGCAGACATAGCAATTTGCACTGTCAGACATGATATACGTTAACTTATGAGG 720
Db 661 ATTGGCAGACATAGCAATTTGCACTGTCAGACATGATATACGTTAACTTATGAGG 720
Qy 721 ACGTATCATAGAAATTTACATATTAACACTTACCGAAAAATTTAAATATTTTACC 780
Db 721 ACGTATCATAGAAATTTACATATTAACACTTACCGAAAAATTTAAATATTTTACC 780
Qy 781 CACATCATCTACACAGAACATATATGCAACGAGATATAGAGAGTATATCTGA 840
Db 781 CACATCATCTACACAGAACATATATGCAACGAGATATAGAGAGTATATCTGA 840
Qy 841 CAACATATCATATTTATTTCTTAACTATCTTACACACAAATTAACGACATATGACACA 900
Db 841 CAACATATCATATTTATTTCTTAACTATCTTACACACAAATTAACGACATATGACACA 900
Qy 901 AATACACACAAAGAGATATGTAATTAATCTCCAAATTTATTTATTTAGTAAAGTAT 960
Db 901 AATACACACAAAGAGATATGTAATTAATCTCCAAATTTATTTATTTAGTAAAGTAT 960
Qy 961 CGAATATCATCTCAATTTTCCAGTGAATAGCAGAAATTTTCAATATATATCTC 1020
Db 961 CGAATATCATCTCAATTTTCCAGTGAATAGCAGAAATTTTCAATATATATCTC 1020
Qy 1021 AATTTTATATCTCTCTGCTCAAGTTTGTCTAAAGAAATTCGAAACATCTTCTGTAG 1080
Db 1021 AATTTTATATCTCTCTGCTCAAGTTTGTCTAAAGAAATTCGAAACATCTTCTGTAG 1080
Qy 1081 TATAGTCTTCTGCTGATATTAAGAACTTGTGATCATGATGAAACATATAGTAA 1140
Db 1081 TATAGTCTTCTGCTGATATTAAGAACTTGTGATCATGATGAAACATATAGTAA 1140
Qy 1141 TGAATGTTCTAAATATATATCCATCAAGATATTTATGCTGACATGAGCTGAT 1200
Db 1141 TGAATGTTCTAAATATATATCCATCAAGATATTTATGCTGACATGAGCTGAT 1200

Qy 1201 AAAGAACATTAATTTGCAATTTGATGGAATATCGCTGAAAAAATTTAGGTAATATGC 1260
Db 1201 AAAGAACATTAATTTGCAATTTGATGGAATATCGCTGAAAAAATTTAGGTAATATGC 1260
Qy 1261 TGTGCTCTGTTATTAATCTTAAAGGATGATGAAATTTAAACCTATATATAGCAAAA 1320
Db 1261 TGTGCTCTGTTATTAATCTTAAAGGATGATGAAATTTAAACCTATATATAGCAAAA 1320
Qy 1321 TCATTCATATAGATCTTCAATCATTAATTTGTCATCATGCTCTGAGATCTAATTTAC 1380
Db 1321 TCATTCATATAGATCTTCAATCATTAATTTGTCATCATGCTCTGAGATCTAATTTAC 1380
Qy 1381 ATGTCACACAGAACATTCACGCAATGAGCTTTTGGACTAGAAAATGAGAAAGGTAGAA 1440
Db 1381 ATGTCACACAGAACATTCACGCAATGAGCTTTTGGACTAGAAAATGAGAAAGGTAGAA 1440
Qy 1441 TATTAATCATTAATATAGATCAACACCTTAAGCTTTGTAAGAAACGATCATGTAAC 1500
Db 1441 TATTAATCATTAATATAGATCAACACCTTAAGCTTTGTAAGAAACGATCATGTAAC 1500
Qy 1501 ACTGATATCATTTAGCAAAAAAATTTATGTAAGAACTACAAATTTGATATCTATTTAA 1560
Db 1501 ACTGATATCATTTAGCAAAAAAATTTATGTAAGAACTACAAATTTGATATCTATTTAA 1560
Qy 1561 TTTATTAATAGATATATATATATATATATATATATATATATATATATATATATAT 1620
Db 1561 TTTATTAATAGATATATATATATATATATATATATATATATATATATATATATAT 1620
Qy 1621 CATCTTTGATATATTTCAACAATGATTAATTAACGCTCAAAAAATATAGATCAACG 1680
Db 1621 CATCTTTGATATATTTCAACAATGATTAATTAACGCTCAAAAAATATAGATCAACG 1680
Qy 1681 TCACACATATATATATATATATATATATATATATATATATATATATATATATAT 1740
Db 1681 TCACACATATATATATATATATATATATATATATATATATATATATATATATAT 1740
Qy 1741 ATATATATATATATATATATATATATATATATATATATATATATATATATATAT 1800
Db 1741 ATATATATATATATATATATATATATATATATATATATATATATATATATATAT 1800
Qy 1801 CCTATCTCTATCTCATATTAACCAATTTGTTATGAGATATTAATTAAGCAATTAAC 1860
Db 1801 CCTATCTCTATCTCATATTAACCAATTTGTTATGAGATATTAATTAAGCAATTAAC 1860
Qy 1861 TTACTTATTTTAAATATATATATATATATATATATATATATATATATATATATAT 1920
Db 1861 TTACTTATTTTAAATATATATATATATATATATATATATATATATATATATATAT 1920
Qy 1921 CATATTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1980
Db 1921 CATATTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1980
Qy 1981 ACCATATATATATATATATATATATATATATATATATATATATATATATATATAT 2040
Db 1981 ACCATATATATATATATATATATATATATATATATATATATATATATATATATAT 2040
Qy 2041 CATATTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2100
Db 2041 CATATTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2100
Qy 2101 TTGCTTATTTTCAAAAAAGTTTAAATTAACATCTGCAACCAATTAATTAATTAATTA 2160
Db 2101 TTGCTTATTTTCAAAAAAGTTTAAATTAACATCTGCAACCAATTAATTAATTAATTA 2160
Qy 2161 TACACGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2220
Db 2161 TACACGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2220
Qy 2221 AATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2280
Db 2221 AATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2280

QY 2281 TCGTATGTAATTTATACAAATTCACAAATTCCTGCAACAACTCAAGTTC 2340
DB 2281 TCGTATGTAATTTATACAAATTCACAAATTCCTGCAACAACTCAAGTTC 2340
QY 2341 TATTAATACAAATTCAGAAATTAATATATCTTACACCAACATAGTTTCAATCTCTA 2400
DB 2341 TATTAATACAAATTCAGAAATTAATATATCTTACACCAACATAGTTTCAATCTCTA 2400
QY 2401 ATACATCATATCTTTTAAAAAGATGACTTACAAATTCATATAATTAATACATG 2460
DB 2401 ATACATCATATCTTTTAAAAAGATGACTTACAAATTCATATAATTAATACATG 2460
QY 2461 CAATGTTTTCAGAAACATTCACAAATTCCTGCAACAAATTCCTGCAACAAATTCCTG 2520
DB 2461 CAATGTTTTCAGAAACATTCACAAATTCCTGCAACAAATTCCTGCAACAAATTCCTG 2520
QY 2521 AGCTTCATTCACAAATTCAGCAATTCATATACCTTCCAAAACGATTAATG 2580
DB 2521 AGCTTCATTCACAAATTCAGCAATTCATATACCTTCCAAAACGATTAATG 2580
QY 2581 AGTTTAAATAGCTCTTCTTCAATTCCTGCAACAAATTCCTGCAACAAATTCCTG 2640
DB 2581 AGTTTAAATAGCTCTTCTTCAATTCCTGCAACAAATTCCTGCAACAAATTCCTG 2640
QY 2641 TTATGACCTTAAACATCATTAATAGCGAATTAAGTATCTATATCTAGTAAATCTTT 2700
DB 2641 TTATGACCTTAAACATCATTAATAGCGAATTAAGTATCTATATCTAGTAAATCTTT 2700
QY 2701 TCGTATACATCTCTTGGAGATTCATTAATTAATCCATTCCTATACATTCGTA 2760
DB 2701 TCGTATACATCTCTTGGAGATTCATTAATTAATCCATTCCTATACATTCGTA 2760
QY 2761 TTGTCTTAAGTACCTGCGCATTCATATAGCTGATGCTATTAAGCTTCACTTAA 2820
DB 2761 TTGTCTTAAGTACCTGCGCATTCATATAGCTGATGCTATTAAGCTTCACTTAA 2820
QY 2821 TACATAGTACAAACCTATTTTTTTTACAGATGCTATTCATTAATTCCTACAAATTC 2880
DB 2821 TACATAGTACAAACCTATTTTTTTTACAGATGCTATTCATTAATTCCTACAAATTC 2880
QY 2881 CGGTATATCTGACAAATTTCTACTAAGACGCTCTTAATTCGAAAGACATTAATCT 2940
DB 2881 CGGTATATCTGACAAATTTCTACTAAGACGCTCTTAATTCGAAAGACATTAATCT 2940
QY 2941 ATATATTAATCTCTCAAAACCTTACAGCTTTTCTTCACTTCAATTCCTTGAAGA 3000
DB 2941 ATATATTAATCTCTCAAAACCTTACAGCTTTTCTTCACTTCAATTCCTTGAAGA 3000
QY 3001 AGACAAATTAATTCACATCTATCTATACAGAAATCTTATATTCCTATATATTT 3060
DB 3001 AGACAAATTAATTCACATCTATCTATACAGAAATCTTATATTCCTATATATTT 3060
QY 3061 TCCCTCTCTTAATTAATCTAGATTCCTTAATTCGCTATGCTATATCTGTTACTTC 3120
DB 3061 TCCCTCTCTTAATTAATCTAGATTCCTTAATTCGCTATGCTATATCTGTTACTTC 3120
QY 3121 ATCTATTAATTAAGTACCAATTCCTGCTCTTAAATTCGCAATGAGGAGTACTCT 3180
DB 3121 ATCTATTAATTAAGTACCAATTCCTGCTCTTAAATTCGCAATGAGGAGTACTCT 3180
QY 3181 ATGAGCAATTAATTAATTAATCTTCTCACTACCAAAATTAATTAACGATTAATG 3240
DB 3181 ATGAGCAATTAATTAATTAATCTTCTCACTACCAAAATTAATTAACGATTAATG 3240
QY 3241 TGGTACATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3300
DB 3241 TGGTACATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3300
QY 3301 TATTAAGCTTAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 3360
DB 3301 TATTAAGCTTAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 3360
QY 3361 CGATGATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3420

DB 3361 CGATGATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3420
QY 3421 GACTATTTCTAATTCCTCAAAATTCCTGATTCGAACTGATTCATTTCTGATATGAC 3480
DB 3421 GACTATTTCTAATTCCTCAAAATTCCTGATTCGAACTGATTCATTTCTGATATGAC 3480
QY 3481 AGACTATAGCTCTCTTTTCAACATTAATTCCTCTCTGTAAGAGCTTTCTAT 3540
DB 3481 AGACTATAGCTCTCTTTTCAACATTAATTCCTCTCTGTAAGAGCTTTCTAT 3540
QY 3541 ATATATCAAGACCAATGAGAGACTTACAGAGTGGCAATTAATTCAGTCCACT 3600
DB 3541 ATATATCAAGACCAATGAGAGACTTACAGAGTGGCAATTAATTCAGTCCACT 3600
QY 3601 AATCATTAATTAACGCAATTAAGATACCTTTCTTTTAACTTTTCCAGTACCTTAATCC 3660
DB 3601 AATCATTAATTAACGCAATTAAGATACCTTTCTTTTAACTTTTCCAGTACCTTAATCC 3660
QY 3661 ATCATATGAGATCTCTTAAACATTAATTCATTAATTAACATGAGGCTTTTCTAT 3720
DB 3661 ATCATATGAGATCTCTTAAACATTAATTCATTAATTAACATGAGGCTTTTCTAT 3720
QY 3721 AGCCATCTTGATGCGGATTAACATCACTGATATTAAGTACATTAATTAATCATCACT 3780
DB 3721 AGCCATCTTGATGCGGATTAACATCACTGATATTAAGTACATTAATTAATCATCACT 3780
QY 3781 TAATATATCTTTTATTAAGATTCGATTAATCACTGATTAATTAATTAATTAATTAAT 3840
DB 3781 TAATATATCTTTTATTAAGATTCGATTAATCACTGATTAATTAATTAATTAATTAAT 3840
QY 3841 AATATCAATCTTTCTGAGCAATTCGAAATTCCTGCAATTAATTAATTAATTAATTAAT 3900
DB 3841 AATATCAATCTTTCTGAGCAATTCGAAATTCCTGCAATTAATTAATTAATTAATTAAT 3900
QY 3901 TGAATATGCAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3960
DB 3901 TGAATATGCAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3960
QY 3961 CTATATGATCAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4020
DB 3961 CTATATGATCAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4020
QY 4021 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4080
DB 4021 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4080
QY 4081 AAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4140
DB 4081 AAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4140
QY 4141 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4200
DB 4141 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4200
QY 4201 TCCATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4260
DB 4201 TCCATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4260
QY 4261 AAGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4320
DB 4261 AAGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4320
QY 4321 GACTATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4380
DB 4321 GACTATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4380
QY 4381 CTTCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4440
DB 4381 CTTCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4440
QY 4441 CCGTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4500

Db 4441 CCCGATATATACATTTATACATATATATGTCAAAACATATCATATGATAGATTTTGT 4500
QY 4501 TATCATGATACCAATTTGGCAAAAGCAAAATATCTTGAATCTGTATGTAAGATTTG 4560
Db 4501 TATCATGATACCAATTTGGCAAAAGCAAAATATCTTGAATCTGTATGTAAGATTTG 4560
QY 4561 GTPACGCTCTTAACTGAGGAGATTTTGAATCCAGAGATGATACATCTATTTTGTTC 4620
Db 4561 GTPACGCTCTTAACTGAGGAGATTTTGAATCCAGAGATGATACATCTATTTTGTTC 4620
QY 4621 TCAATCAGCAAAAGCAGCAAAAAGATAGACCTTGATCAACAAAGTTAGCTATTTGAT 4680
Db 4621 TCAATCAGCAAAAGCAGCAAAAAGATAGACCTTGATCAACAAAGTTAGCTATTTGAT 4680
QY 4681 CAGAACTTACATTTATTTATACCATCATATATTTTATCAACAGTAATCCACAAATGA 4740
Db 4681 CAGAACTTACATTTATTTATACCATCATATATTTTATCAACAGTAATCCACAAATGA 4740
QY 4741 CTATATGATC 4750
Db 4741 CTATATGATC 4750

RESULT 3
AX685249 1416 bp DNA linear PAT 29-MAR-2003
LOCUS AX685249 Sequence 56 from Patent WO0206652.
DEFINITION AX685249
ACCESSION AX685249
VERSION AX685249.1 GI:29371591
KEYWORDS
SOURCE Ehrlichia ruminantium (heartwater rickettsia)
ORGANISM Ehrlichia ruminantium
Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
Anaplasmataceae; Ehrlichia.
REFERENCE 1
AUTHORS Barbet, A.F., Whicmire, W.M., Kamper, S.M., Simbl, B.H., Ganta, R.R.,
Worland, A.L., Weng, D.M., Meguire, T.C. and Khan, S.M.
TITLE Ehrlichia ruminantium polypeptides, antigens, polynucleotides, and
methods of use
JOURNAL Patent: WO 0206652-A 56 29-AUG-2002;
UNIVERSITY OF FLORIDA (US)
FEATURES
source location/Qualifiers
1..1416
/organism="Ehrlichia ruminantium"
/mol_type="unassigned DNA"
/db_xref="taxon:779"
misc_feature 1..1416
/note="Complement to SEQ ID NO:52, nucleotides 2467-3882
Hypothetical nitrogen assimilation regulatory protein
Product = 19HMRP3"

ORIGIN
Query Match 29.8%; Score 1416; DB 6; Length 1416;
Best Local Similarity 100.0%; Pred. No. 1,38-169;
Matches 1416; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2467 TTAATCAGAAATATACCAATTCAGAAATTTTCAATTTAGCTGTATGCTGAACGCTTC 2526
Db 1 TTAATCAGAAATATACCAATTCAGAAATTTTCAATTTAGCTGTATGCTGAACGCTTC 60
QY 2527 CATTCGCAAAATTCAGAGCTTCGTGATACATTAATCTCCAAACAGATACTGAGTTT 2586
Db 61 CATTCGCAAAATTCAGAGCTTCGTGATACATTAATCTCCAAACAGATACTGAGTTT 120
QY 2587 TAAATATCTCTTTCAAAATTTCTTCAAGAGCTTTAGCTAATGCTACAGAAATTAATCTTTAGC 2646
Db 121 TAAATATCTCTTTCAAAATTTCTTCAAGAGCTTTAGCTAATGCTACAGAAATTAATCTTTAGC 180
QY 2647 ACTTAAACATCATTAATAGCGAATTAATAGATCTATATCTAGCTGTAATCTTTTGTCTGT 2706
Db 181 ACTTAAACATCATTAATAGCGAATTAATAGATCTATATCTAGCTGTAATCTTTTGTCTGT 240
QY 2707 AATCATCTCTTGAAGATTTTCAATTAATTCATTCATTAATAGCTAATCTGT 2766

Db 241 AATCATCTCTTGAAGATTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 300
QY 2767 TAAATATCTCTTTCAAAATTTCTTCAAGAGCTTTAGCTAATGCTACAGAAATTAATCTTTAGC 2826
Db 301 TAAATATCTCTTTCAAAATTTCTTCAAGAGCTTTAGCTAATGCTACAGAAATTAATCTTTAGC 360
QY 2827 AGTACACAAACCTATTTTCTTCAAGAGCTTTAGCTAATGCTACAGAAATTAATCTTTAGC 2886
Db 361 AGTACACAAACCTATTTTCTTCAAGAGCTTTAGCTAATGCTACAGAAATTAATCTTTAGC 420
QY 2887 ATCTGTCAATTAATTTCTTCAAGAGCTTTAGCTAATGCTACAGAAATTAATCTTTAGC 2946
Db 421 ATCTGTCAATTAATTTCTTCAAGAGCTTTAGCTAATGCTACAGAAATTAATCTTTAGC 480
QY 2947 TAAATCTCTCAAAACCTATTTTCTTCAAGAGCTTTAGCTAATGCTACAGAAATTAATCTTTAGC 3006
Db 481 TAAATCTCTCAAAACCTATTTTCTTCAAGAGCTTTAGCTAATGCTACAGAAATTAATCTTTAGC 540
QY 3007 AATATCTCTCAAAACCTATTTTCTTCAAGAGCTTTAGCTAATGCTACAGAAATTAATCTTTAGC 3066
Db 541 AATATCTCTCAAAACCTATTTTCTTCAAGAGCTTTAGCTAATGCTACAGAAATTAATCTTTAGC 600
QY 3067 CTGTATTAATCTGTATTAATCTTTAGCTAATGCTACAGAAATTAATCTTTAGC 3126
Db 601 CTGTATTAATCTGTATTAATCTTTAGCTAATGCTACAGAAATTAATCTTTAGC 660
QY 3127 AATTAACGTAACGTAATTTCTTCAAGAGCTTTAGCTAATGCTACAGAAATTAATCTTTAGC 3186
Db 661 AATTAACGTAACGTAATTTCTTCAAGAGCTTTAGCTAATGCTACAGAAATTAATCTTTAGC 720
QY 3187 CAATATATTAATTTCTTCTTCAAGAGCTTTAGCTAATGCTACAGAAATTAATCTTTAGC 3246
Db 721 CAATATATTAATTTCTTCTTCAAGAGCTTTAGCTAATGCTACAGAAATTAATCTTTAGC 780
QY 3247 CATATGATGATATGATATTAATGATATTAATGATATTAATGATATTAATGATATTAATGATATTA 3306
Db 781 CATATGATGATATGATATTAATGATATTAATGATATTAATGATATTAATGATATTAATGATATTA 840
QY 3307 CCTAGCACTACTCTTCTTCAAGAGCTTTAGCTAATGCTACAGAAATTAATCTTTAGC 3366
Db 841 CCTAGCACTACTCTTCTTCAAGAGCTTTAGCTAATGCTACAGAAATTAATCTTTAGC 900
QY 3367 AGTATGCTCTTATTAATCAATTTCTTCAAGAGCTTTAGCTAATGCTACAGAAATTAATCTTTAGC 3426
Db 901 AGTATGCTCTTATTAATCAATTTCTTCAAGAGCTTTAGCTAATGCTACAGAAATTAATCTTTAGC 960
QY 3427 TTAATATCTCTTCAAAATTTCTTCAAGAGCTTTAGCTAATGCTACAGAAATTAATCTTTAGC 3486
Db 961 TTAATATCTCTTCAAAATTTCTTCAAGAGCTTTAGCTAATGCTACAGAAATTAATCTTTAGC 1020
QY 3487 TATAGCTCTTCTTCAAAATTTCTTCAAGAGCTTTAGCTAATGCTACAGAAATTAATCTTTAGC 3546
Db 1021 TATAGCTCTTCTTCAAAATTTCTTCAAGAGCTTTAGCTAATGCTACAGAAATTAATCTTTAGC 1080
QY 3547 ATAAGACCCATATGAGAGACTTTACAGAGGAGGATTAATGAGGAGGATTAATGAGGAGGATTAAT 3606
Db 1081 ATAAGACCCATATGAGAGACTTTACAGAGGAGGATTAATGAGGAGGATTAATGAGGAGGATTAAT 1140
QY 3607 AATTAACGGAATTAAGATTAATCTTTCAAGAGCTTTAGCTAATGCTACAGAAATTAATCTTTAGC 3666
Db 1141 AATTAACGGAATTAAGATTAATCTTTCAAGAGCTTTAGCTAATGCTACAGAAATTAATCTTTAGC 1200
QY 3667 ATGAGATCTCTTCAAAATTTCTTCAAGAGCTTTAGCTAATGCTACAGAAATTAATCTTTAGC 3726
Db 1201 ATGAGATCTCTTCAAAATTTCTTCAAGAGCTTTAGCTAATGCTACAGAAATTAATCTTTAGC 1260
QY 3727 CTGTATGCGGATTAACCAATCTGCTAATTTAGTACATTAATTAATCTTTAGC 3786
Db 1261 CTGTATGCGGATTAACCAATCTGCTAATTTAGTACATTAATTAATCTTTAGC 1320
QY 3787 ATCTTTATTAAGTTTCTGATATCAATCTTCAATCAATCAATCAATCAATCAATCAATCAATCAAT 3846

DB 1321 ATCTTTATGATTTGTGATATCACTTCATCAACAACTAATTAATTGAGAAATATA 1380

QY 3847 CAATCTTCCTGGACATTTCAAAATCCTGGACAT 3882

DB 1381 CAATCTTCCTGGACATTTCAAAATCCTGGACAT 1416

RESULT 4
AX685247 984 bp DNA linear PAT 29-MAR-2003
LOCUS AX685247
DEFINITION Sequence 54 from Patent WO02066652.
ACCESSION AX685247 GI:29371589
VERSION AX685247.1
KEYWORDS
SOURCE
ORGANISM *Escherichia ruminantium* (heartwater rickettsia)
Escherichia ruminantium
Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
Anaplasmataceae; Ehrlichia.

REFERENCE
AUTHORS Barbet, A.F., Whitmore, W.M., Kamper, S.M., Simbi, B.H., Ganta, R.R.,
Moreland, A.L., Wang, D.M., McGuire, T.C. and Mahan, S.M.
TITLE Ehrlichia ruminantium polypeptides, antigens, polynucleotides, and
methods of use
JOURNAL Patent: WO 02066652-A 54 29-AUG-2002;
UNIVERSITY OF FLORIDA (US)

FEATURES
Source
1..984
/organism="Ehrlichia ruminantium"
/mol_type="unassigned DNA"
/db_xref="taxon:779"
1..984
/note="unassigned protein product; Corresponds to SEQ ID
NO:52, nucleotides 632-1615 Hypothetical glycerol
3-phosphate dehydrogenase Product = 19HMOF2"
/codon_start=1
/protein_id="CAD86555.1"
/db_xref="GI:29371590"
/db_xref="EMBL:CAD86555"
/translation="MKISILGAGSGRTAIALSAGISVNLGRDHNTHTINTYK
NKLPTFHEPNTIYATSNIDEVLSNNITCITLPTPOQLTCTQLOHOKNKP
ILICSGEITSLKPESEIAEILQVNTIFLSPSPAKETAEHLPGIVAGNKL
GESLITISNDLKIYHQDIQVQALAKNIATACGIAAGNLNNNAVATITIG
MWEIKLYIANHSIDILHTLIGPSCLGILITCTEBSRNAFGLEIGKGNITLLID
HMLKVEGSTYKPLISAKUNVELPICISIVNLHENISLDXAINILS"

ORIGIN
Query Match 20.7%; Score 984; DB 6; Length 984;
Best Local Similarity 100.0%; Pred. No. 4,46-115;
Matches 984; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 632 ATGAAATCAGATTTTGGTGCAGATCAATTTGGCAGAGCAATAGCAATGCACTGCA 691

DB 1 ATGAAATCAGATTTTGGTGCAGATCAATTTGGCAGAGCAATAGCAATGCACTGCA 60

QY 692 GCAATGCTATTCAGTTACTTATGAGGAGCTGATCTAGAAATATTACATATAAC 751

DB 61 GCAATGCTATTCAGTTACTTATGAGGAGCTGATCTAGAAATATTACATATAAC 120

QY 752 ACTTACCGAAAAATTTAAATATTACCACATATCTTACCGACAAATATATGCA 811

DB 121 ACTTACCGAAAAATTTAAATATTACCACATATCTTACCGACAAATATATGCA 180

QY 812 ACCAGCAATATAGCAGAGTATTTATCTGCAACAATACATGATATTCTTAATTCCT 871

DB 181 ACCAGCAATATAGCAGAGTATTTATCTGCAACAATACATGATATTCTTAATTCCT 240

QY 872 ACACAAATATGACGACATATGATACAAATATACAAACAAACAGCATATGTTAAAT 931

DB 241 ACACAAATATGACGACATATGATACAAATATACAAACAAACAGCATATGTTAAAT 300

QY 932 ACTCCAAATTTATTTGTGTAAGGTATGAAATTAATCACTCAATTTCCAGTGAA 991

DB 301 ACTCCAAATTTATTTGTGTAAGGTATGAAATTAATCACTCAATTTCCAGTGAA 360

QY 992 ATAGCAGAAATTTTCAATATATCAATTTTATACCTGCTGGCCAGATTGCT 1051

DB 361 ATAGCAGAAATTTTCAATATATCAATTTTATACCTGCTGGCCAGATTGCT 420

QY 1052 AAAGAAATTCAGAAATCTTCTGTGTATGATGACTGCTGTGTATTAAGAACTT 1111

DB 421 AAAGAAATTCAGAAATCTTCTGTGTATGATGACTGCTGTGTATTAAGAACTT 480

QY 1112 GGTGAATCTTATGTAAGAAATTAAGTAAGTGTCTTAAAAATATATACATCAAGAT 1171

DB 481 GGTGAATCTTATGTAAGAAATTAAGTAAGTGTCTTAAAAATATATACATCAAGAT 540

QY 1172 ATTATAGGTATCAGATTGGAGCTGCATTAAGAAATATATGCAATTCATGTTGAATA 1231

DB 541 ATTATAGGTATCAGATTGGAGCTGCATTAAGAAATATATGCAATTCATGTTGAATA 600

QY 1232 ATGCTGCAAAAAATTTAGTATATATGCTGTGCTACTCTTATATCAAGGCAATGAT 1291

DB 601 ATGCTGCAAAAAATTTAGTATATATGCTGTGCTACTCTTATATCAAGGCAATGAT 660

QY 1292 GAAATTAACACATATATATGCAAAAAATCATTCATATGATCTTCATATCAATTAATGCT 1351

DB 661 GAAATTAACACATATATATGCAAAAAATCATTCATATGATCTTCATATCAATTAATGCT 720

QY 1352 CCATCATGCTTGGAGATTTATTTATACATGATATACAGAAACATTCACGCAATGCT 1411

DB 721 CCATCATGCTTGGAGATTTATTTATATATATATATATATATATATATATATATGCT 780

QY 1412 TTGGACTAGAAATAGAAAGAGTAAATATATATATATATATATATATATATATGCT 1471

DB 781 TTGGACTAGAAATAGAAAGAGTAAATATATATATATATATATATATATATATGCT 840

QY 1472 CTGTGTAGAGAACCAAGTATCTGTAACACCTGATATCTTTGCAAAAAATCTTAATGTA 1531

DB 841 CTGTGTAGAGAACCAAGTATCTGTAACACCTGATATCTTTGCAAAAAATCTTAATGTA 900

QY 1532 GAATCAATTTGCAATATCTATTTATCAATTTATATCAATGAGAAATATCACTAGATAA 1591

DB 901 GAATCAATTTGCAATATCTATTTATCAATTTATATCAATGAGAAATATCACTAGATAA 960

QY 1592 GCCATTCAAACATATATCTTATG 1615

DB 961 GCCATTCAAACATATATCTTATG 984

RESULT 5
AX685250 597 bp DNA linear PAT 29-MAR-2003
LOCUS AX685250
DEFINITION Sequence 57 from Patent WO02066652.
ACCESSION AX685250
VERSION AX685250.1 GI:29371592
KEYWORDS
SOURCE
ORGANISM *Escherichia ruminantium* (heartwater rickettsia)
Escherichia ruminantium
Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
Anaplasmataceae; Ehrlichia.

REFERENCE
AUTHORS Barbet, A.F., Whitmore, W.M., Kamper, S.M., Simbi, B.H., Ganta, R.R.,
Moreland, A.L., Wang, D.M., McGuire, T.C. and Mahan, S.M.
TITLE Ehrlichia ruminantium polypeptides, antigens, polynucleotides, and
methods of use
JOURNAL Patent: WO 02066652-A 57 29-AUG-2002;
UNIVERSITY OF FLORIDA (US)

FEATURES
Source
1..597
/organism="Ehrlichia ruminantium"
/mol_type="unassigned DNA"
/db_xref="taxon:779"
1..597
/note="Complement to SEQ ID NO:52, nucleotides 4154-4750
Hypothetical cell division protein ftsQ Product =

ORIGIN

19HMORF41"

Query Match 12.6%; Score 597; DB 6; Length 597;
 Best Local Similarity 100.0%; Pred. No. 3.4e-66;
 Matches 597; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4154 TTATTCATATATATTAACCTTCCGGAACAGCAGCATGTATAGTATTTCCATACATAAA 4213
 Db 1 TTATTCATATATATTAACCTTCCGGAACAGCAGCATGTATAGTATTTCCATACATAAA 60
 QY 4214 CTCACTTGAAGCTTTATAGATATTAATTAATTCAGCCAGCAGCATTAAGATTCATATT 4273
 Db 61 CTCACTTGAAGCTTTATAGATATTAATTAATTCAGCCAGCAGCATTAAGATTCATATT 120
 QY 4274 AGGTATTTATATTAATTAACCACTAGACAAAATTAATCCACCATGACTATCCAACTA 4333
 Db 121 AGGTATTTATATTAATTAACCACTAGACAAAATTAATCCACCATGACTATCCAACTA 180
 QY 4334 TGTGATAGATGAACCAATACCAACCAAGAGTATTAATGACTACTCTCACTATATAA 4393
 Db 181 TGTGATAGATGAACCAATACCAACCAAGAGTATTAATGACTACTCTCACTATATAA 240
 QY 4394 ATCTAATAGTGAAGCAGCAGCATCTCATGATGTGAAGTCAATCATCCCGATATTTACA 4453
 Db 241 ATCTAATAGTGAAGCAGCAGCATCTCATGATGTGAAGTCAATCATCCCGATATTTACA 300
 QY 4454 ATTATCAACATATATATGTCACAAATCATCAATGATAGAAATTTGTTATCATGATACA 4513
 Db 301 ATTATCAACATATATATGTCACAAATCATCAATGATAGAAATTTGTTATCATGATACA 360
 QY 4514 ATTTCGAAAAGCGAATATTTCTTGAACGTGATCTGTAAAGTATTTGGTAACAGCTCTT 4573
 Db 361 ATTTCGAAAAGCGAATATTTCTTGAACGTGATCTGTAAAGTATTTGGTAACAGCTCTT 420
 QY 4574 AACTGAGGCAATTTTATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4633
 Db 421 AACTGAGGCAATTTTATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 QY 4634 AGGCACAAAAGATAGACCTTGATCAACCAAGTTAGCTATTTATCATGAACTTACATA 4693
 Db 481 AGGCACAAAAGATAGACCTTGATCAACCAAGTTAGCTATTTATCATGAACTTACATA 540
 QY 4694 TTGATTAACATCAATGATATTTTATCAACAGTAATTCACAACTTGCATTAATGATC 4750
 Db 541 TTGATTAACATCAATGATATTTTATCAACAGTAATTCACAACTTGCATTAATGATC 597

RESULT 6
 LOCUS AX685246 371 bp DNA linear PAT 29-MAR-2003
 DEFINITION Sequence 53 from Patent WO02066652.
 ACCESSION AX685246
 VERSION AX685246.1 GI:29371588
 KEYWORDS
 SOURCE
 ORGANISM
 Ehrlichia ruminantium (heartwater rickettsia)
 Ehrlichia ruminantium
 Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 Anaplasmataceae; Ehrlichia.

REFERENCE
 AUTHORS 1 Barbet, A.F., Whitmore, W.M., Kamper, S.W., Simbi, B.H., Ganta, R.R.,
 Moreland, A.L., Mwambi, D.M., McGuire, T.C., and Mahan, S.M.
 TITLE Ehrlichia ruminantium polypeptides, antigens, polynucleotides, and
 methods of use
 JOURNAL Patent: WO 02066652-A 53 29-AUG-2002;
 UNIVERSITY OF FLORIDA (US)
 FEATURES
 source
 1..371
 /location/Qualifiers
 /organism="Ehrlichia ruminantium"
 /mol_type="unassigned DNA"
 /db_xref="taxon:779"
 1..371
 /note="Complement to SEQ ID NO:52, nucleotides <1_371

ORIGIN

Hypothetical integrase/recombinase Product = 19HMORF11"

Query Match 7.8%; Score 371; DB 6; Length 371;
 Best Local Similarity 100.0%; Pred. No. 1.3e-37;
 Matches 371; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATCCGCTTAATTAATTAAGTATTTTAAAGATGCGCCAGTAAATTAATTAAGTACT 60
 Db 1 GATCCGCTTAATTAATTAAGTATTTTAAAGATGCGCCAGTAAATTAATTAAGTACT 60
 QY 61 GCTTACTGCTATATATAGTACATACGTATGATATTAACGAAAAAATTTCTCAAGCAGAC 120
 Db 61 GCTTACTGCTATATATAGTACATACGTATGATATTAACGAAAAAATTTCTCAAGCAGAC 120
 QY 121 AGTACCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
 Db 121 AGTACCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
 QY 181 CTCAATCTCCATTTTAACATGACACAGATTTATTAATGTAACAGAACACCTGTACTC 240
 Db 181 CTCAATCTCCATTTTAACATGACACAGATTTATTAATGTAACAGAACACCTGTACTC 240
 QY 241 TTATGCAAGAACTTATTAACCTATCAAGGTCTCTTACATTAAGAACAACTGATTTAAA 300
 Db 241 TTATGCAAGAACTTATTAACCTATCAAGGTCTCTTACATTAAGAACAACTGATTTAAA 300
 QY 301 GAATACCTCTCTGTTGTCATCATATTTTCCATATTTTCAATATTAATCATATAGAGTCT 360
 Db 301 GAATACCTCTCTGTTGTCATCATATTTTCCATATTTTCAATATTAATCATATAGAGTCT 360
 QY 361 TTATTCGACAT 371
 Db 361 TTATTCGACAT 371

RESULT 7

REPX03/c

LOCUS REPX03 279110 bp DNA linear BCT 02-SEP-2002
 DEFINITION Rickettsia prowazekii strain Madrid E, complete genome; segment 3/4.

ACCESSION AJ235272 AJ235269
 VERSION AJ235272.1 GI:3861033
 KEYWORDS complete genome.
 SOURCE Rickettsia prowazekii
 ORGANISM Rickettsia prowazekii
 Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 Rickettsiaceae; Rickettsiae; typhus group.

REFERENCE
 AUTHORS 1 Andersson, S.G., Zomorodipour, A., Andersson, J.O.,
 Sichterich-Ponten, T., Almaraz, U.C., Podowski, R.M., Naslund, A.K.,
 Eriksson, A.S., Winkler, H.H. and Kurland, C.G.
 TITLE The genome sequence of Rickettsia prowazekii and the origin of
 mitochondria
 JOURNAL Nature 396 (6707), 133-140 (1998)

JOURNAL MEDLINE
 PUBMED 9823893
 NATURE 396 (6707), 133-140 (1998)

REFERENCE
 AUTHORS 2 (bases 1 to 279110)
 TITLE Direct Submission
 JOURNAL Submitted (11-NOV-1998) S.G.E. Andersson,
 Siv. Andersson-Olsson, U. Se, Dept. of Molecular Biology, University
 of Uppsala, Husargatan 3, Uppsala, S-751 24, SWEDEN
 FEATURES
 source
 1..279110
 /organism="Rickettsia prowazekii"
 /mol_type="genomic DNA"
 /strain="Madrid E"
 /db_xref="taxon:782"
 63..1046
 /gene="Rp479"
 63..1046
 /gene="Rp479"

CDS

gene
 63..1046
 /gene="Rp479"
 63..1046
 /gene="Rp479"


```

/protein_id="CAI4942.1"
/db_xref="GI:3861042"
/db_xref="GOA:Q92D57"
/db_xref="Swiss-Prot:Q92D57"
/translation="MYLSEIIGLNTLRLQLRLSISIDPYKDYKYNQGYIKYLE
TLEFIPILYICILNYITLKDYNGTSSKVTNDEYIINLPKIKNNKSIYEE
VBPILYKNNKNEFVIDTNGVSNOKSVPLENLTKNLVSNTKHPISIVNE
SEI EKONEVILPEIILKKYFADNLAYPNLFIPEGTVLILFEPVTELLERSTIYLV
YGLNLTETKTSIOTSRILVMPSSGIPILQPVIIILILBELSILIDLMFTFLVFLV
ALLQINKSLSHIT"
complement(9681..10349)
/gene="RP490"
/complement(9681..10349)
/gene="RP490"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="CAI4943.1"
/db_xref="GI:3861043"

Query Match      7.5%: Score 357.6; DB 1: Length 279110;
Best Local Similarity 53.1%; Pred. No. 5,9e-37;
Matches 1021; Conservative 0; Mismatches 849; Indels 54; Gaps 10;

QY 2397 TATAATACATCATCTCTTTTAAAAAGATGACTTCACATTCATACATTAATATTATA 2456
Db 110538 TAAAGAGAAATATATATGATATGATTTAGCATTTGCTTCTCATATCTTCTT 110479
QY 2457 CATGCATGTTTATTCAGAAACATTCACATCCAGAAATTTTCATTTACGGTGAAG 2516
Db 110478 CATTTATTCATTTGTTGGGGGATATGTTAGCTTAAATTTTATGTTTCATGTTAAAG 110419
QY 2517 CTGAAGCTTCATTCACAAATTCAGACGTTGCTGATACATTAACCTCCAAAAGAGATA 2576
Db 110418 CCGATCTTCACATCCAGAAATGAGAAAGTTTGAATATATTATTTAAACGACTCA 110359
QY 2577 ACTGAGTTTAAATACGTTCTTCAAAATCTTCAGAGGTTTACGTAATGTCAGAGAA 2636
Db 110358 TTGTGCTGATAGATATGATTCGTTCAAAAACCTTCTCAGCTTCTTAAGGGGATAGATA 110299
QY 2637 TAACTTAGACATTAACATCATTT-----AATAGCGATATAGATATATCTACTG 2690
Db 110298 ACATATCAAGCATCTTCAAGTTAGTAAAGTTAGACATATTTGCTAATATTTCTGAAG 110239
QY 2691 GTAAATCTTTGCTGATATCTCTCTAGGA-----GATTCATATTTAAAA 2738
Db 110238 GAATCATATAGGTTTATATATTTCAATTTACCTGATTAACGATTCATATTAAG 110179
QY 2739 TCCATTCATACATTCATGATTTGTTAGTTTACCTGCGCATCATATGACTGATG 2798
Db 110178 TCCATTCACAGGTTACGTAATGTTCTAATATTCAGGCCATTTGATGATGAAAGG 110119
QY 2799 CTATTAAGGTTTATCTCTTAATACATGATGATACAAACCTATTTTTCAGAGTCAT 2858
Db 110118 CTGCAATAGGTTTATGAGAAAGTGAACGTTCTTTTAAACCTGAAAATTTTGAAGTTGCT 110059
QY 2859 TCAATTAATACCTACACAAATTCGGATATCTGTACAAATTTTCTACATAAGACGCTACTG 2918
Db 110058 TACCAAAATATTTAATCAATAGGATATATCTTCTTCTTCTCATATTAATGAAGGACTT 109939
QY 2919 TAAATGAGAGCAATTAATCTATTAATATATCTTCACAAAACCTACAGGCTTTTACTT 2978
Db 109938 TTAAGAAAAAACAATTAAGCGAATATATGATCTTCTAGAAATTTCCGTTATTTACTT 109339
QY 2979 CACTTCATATCTTTGAGAAAGACAAATATTCACATCTATACATAAGAGATCT 3038
Db 109338 CACTTCGATATTTTTCAGCTGATCTGATATATTTTATATACCTTTAGCTTTTTC 109879
QY 3039 TACTATTTTCCATATATTTTCCCTCTGTATATATCTGAGTAATCTTAAATGCGAT 3098
Db 109878 CGGAAGGTTTGTAA---ATAGTTGATCTTTAAGAAATTTATATATTTTACCTGATAG 109822
QY 3099 CGATGTAATCTGTTACTTATCTATATAATAGCATACATGATTTGCTTCTATATA 3158

```

```

Db 109821 GAATAGAAATATATCTGATCTCATCTATATATTAAGTAAACCTTATTTGCAAAATCTA--- 109765
QY 3159 TTCCAAATGAGAGAGTACTCTATAGACAAATATATTATTTCTTCCCTACATCAATA 3218
Db 109764 -----AGATAGTAGAGAGGTTTATATATATTTCTTCCGCTTTTCCGATTCGCAAT 109715
QY 3219 TATTACCAAGTAATTTATTTACTGATGATAGATAGATGATAGATATGAT ATAAATGAGATA 3277
Db 109714 AATCTTGATTAATTTTCTGATGATAGATAGAGGTTGACATAAAAATATGAATGATTA 109655
QY 3278 TCATACCCCTTGATTTTATGATATAGCTTACCAACTACTCTTCTTCCAAACCTGGC 3337
Db 109654 TTAAGCTCTTGAATTTGTTTATGATTAATTAACCTTCTGCAAGTTCTTTGGCCGCTACGACT 109595
QY 3338 GAACCAATAGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3397
Db 109594 TTACCATGAATCAATTAATACGATCTGATGATGATGATGATGATGATGATGATGATGAT 109535
QY 3398 TTACGATATACAGGAGGATTCACGATATTTCA---TAATCCGAATGCTGATTTGAC 3454
Db 109534 TTTAAGTTACCGAGATCTCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 109475
QY 3455 TCATCATTTTCTGATGATATCTACAGACTCTATAGCTCTTTTACAACTTAAAT 3514
Db 109474 TCTATGTTTTCGCTTTTAACTTTGTTACTTCGCAAGCTCTTTTAAATTAATTAAT 109415
QY 3515 CTGCTGCTGTAAGAGGCTTTCTATATATATCAATACAGACCCATATGACAGACTTACA 3574
Db 109414 TTAATATTTAATATGTTTCTCAATATATGATGATGATGATGATGATGATGATGATGAT 109355
QY 3575 GAGTGGCAATTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3634
Db 109354 GCTGTTCTATAGTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 109295
QY 3635 TTAAGCTTTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3694
Db 109294 TTAATATCTTAAATACCTCAAGCCGATATTTGCTGATGATGATGATGATGATGATGAT 109235
QY 3695 AATCAATCAATGAGCTCTTTTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 3754
Db 109234 ATCACTGAGAGACCGGTTTCTGGAATATTTTAAAGAGTTGATGATGATGATGATGATGAT 109175
QY 3755 AATTTAGGATATATATATATATATATATATATATATATATATATATATATATATATATATAT 3814
Db 109174 ATCTTAGATTTAAACCTTCACTTTTAAATTTGACAAATGATGATGATGATGATGATGAT 109115
QY 3815 TCATCATCAACATTAATATCTTCAAGATATATATATATATATATATATATATATATATATATAT 3874
Db 109114 TCAATGCTATATATTAATTAATCAATTTGATGATGATGATGATGATGATGATGATGATGAT 109058
QY 3875 TGTGCAATATTTACTTACATATATATATATATATATATATATATATATATATATATATATAT 3934
Db 109057 AAAAAACCAATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 108998
QY 3935 TACACCTGATATATATATATATATATATATATATATATATATATATATATATATATATATAT 3994
Db 108997 TTTTGTGCAATATATATATATATATATATATATATATATATATATATATATATATATATAT 108940
QY 3995 TTTTACTTATTAATTTAAACATATATATATATATATATATATATATATATATATATATATATAT 4054
Db 108939 TATTAGTTGTTTATTAATATATATATATATATATATATATATATATATATATATATATATAT 108881
QY 4055 AAGCATAGGTTTCTATATATATATATATATATATATATATATATATATATATATATATATAT 4114
Db 108880 -----AGCTTATTAATATATATATATATATATATATATATATATATATATATATATAT 108828
QY 4115 ATCACTTTGTAACCTATATATATATATATATATATATATATATATATATATATATATATATAT 4174
Db 108827 AGTGTGTTTGTGATATATATATATATATATATATATATATATATATATATATATATATAT 108768
QY 4175 CCC---TGGAACAGGATGCTATAGTTTCCATGATGATGATGATGATGATGATGATGATGATGAT 4231

```


Db 108767 CACTGATGTAATATAGTATGATATGATACGCAATAAAGCTCTCACTAGTAGTAA 108708
QY 4232 GATATTCATTAATTCACGGCAAGCATTAATCATTTATAGTATTTATTTAA 4231
Db 108707 TAAATTTTAAATATATGCTTAAACATCTTTTAAATTTATTTGCAATTTATTTGCTTA 108648
QY 4292 ACCA 4295
Db 108647 ACCA 108644

RESULT 8
AE008640/c 12477 bp DNA linear BCT 14-SEP-2001
LOCUS Rickettsia conorii Malish 7, section 72 of 114 of the complete
DEFINITION genome.
ACCESSION AE008640 AE006914
VERSION AE008640.1 GI:15619951
KEYWORDS
SOURCE Rickettsia conorii
ORGANISM Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
Rickettsia conorii
Rickettsiaceae; Rickettsiae; Rickettsia; spotted fever group.
1 (sites)
Ogata,H., Audic,S., Barbe,V., Artiguenave,F., Fournier,P.E.,
Raoult,D. and Claverie,J.M.
TITLE Selfish DNA in protein-coding genes of Rickettsia
JOURNAL Science 290 (5490), 347-350 (2000)
MEDLINE 20485642
PUBMED 11030655
REFERENCE 2 (sites)
Ogata,H., Audic,S. and Claverie,J.-M.
TITLE Selfish DNA and the origin of genes
JOURNAL Science 291 (5502), 252-253 (2001)
REFERENCE 3 (bases 1 to 12477)
Ogata,H., Audic,S., Renesto-Audiffren,P., Fournier,P.-E., Barbe,V.,
Samson,D., Roux,V., Cosserat,P., Weissenbach,J., Claverie,J.-M. and
Raoult,D.
TITLE Mechanisms of evolution in Rickettsia conorii and R. prowazekii
JOURNAL Science 293 (5537), 2093-2098 (2001)
MEDLINE 21442074
PUBMED 11557893
REFERENCE 4 (bases 1 to 12477)
Ogata,H., Audic,S., Renesto-Audiffren,P., Fournier,P.-E., Barbe,V.,
Samson,D., Roux,V., Cosserat,P., Weissenbach,J., Claverie,J.-M. and
Raoult,D.
TITLE Direct Submission
JOURNAL Submitted (26-APR-2001) Unite des Rickettsies, Faculte de Medecine,
CNR UMR 6020, Universite de la Mediterranee, 27 boulevard Jean
Moulin, Marseille Cedex 05 13385, France
COMMENT A public version of R. conorii genome database is accessible at
http://igs-server.cnr-mrs.fr/. The database intends to provide
updated data. Annotation of the genome is an ongoing task whose
goal is to make the genome sequence more useful. Comments to the
authors are appreciated.
FEATURES
source location/Qualifiers
1..12477
/organism="Rickettsia conorii"
/mol_type="genomic DNA"
/strain="Malish 7"
/db_xref="taxon:781"
complement(169..267)
/note="RBP0, repeated element"
465..636
/note="RBP06, A and T rich repeated element"
664..2088
/gene="ntrX"
/note="synonym: RC0849"
664..2088
/gene="ntrX"
/codon_start=1
/transl_table=11
/product="nitrogen assimilation regulatory protein ntrX"

gene
CDS
/protein_id="AA103387.1"
/db_xref="GI:15619952"
/translation="MSQDVLIVYDDEBSIRNLIANLKQEGNPKYANSTQALKLS
KEVSAVAVLIDWLQSGEIDGLILEIKRYPLEPVIIISGRTLETAVNAIKMGAYD
YIEKFPNDKLVILLRACEVRLKRENDLVKSDYIDKTELVEGSEVTLAKMAAKA
ATSSCRIMHGVKSGSELAARLIHKOSRVANNPFIIEPTCMTEKINQIEFGELEK
QANAKPTLIEFANNGTLYIDEVSNIPPIQVYKLFKLDQITTKPCGNIKIDIKI
TGTSKTIQDEVNNGKLEBLYLRINVSILKPSLYEKEDIPLYKFPKQSKPSGL
KRSFADETIALQSYEMRGNROLNVEMLINPLITGNEELIKPYMTSEILAN
SANITLEDSPDMSMPLEAREBVERQYLSQMSRPNINIKTSFPMGERSALHRK
LKULSLHPTPRINDEYEANA"
2057..3826
/gene="RC0850"
2057..3826
/gene="RC0850"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AA103388.1"
/db_xref="GI:15619953"
/translation="MTKMRKQMLKVIISITTYLLSGSESTRDANGHLTDSQSTVI
RVYIISQSKNLKMLKEKFGSLNGVLIQVILNEDLSGIDLSCELTADPAGSN
LKKAILTNALIOESNPDSDVINKIGSHDPGSGI FHNITLQNTNPFQDPAFK
TTIINPENSRSKSHLMNNTIDGVNPKALQNNSPKNTITSIPGCTLEKSI
KTNFPNNPFSNNLSQITLIVITIKDSPTOSIENYFNVSQNSGCFSTAFQDS
TLQNISLTQCDLQNSTISSVLKFRINNALINNSLNDKNTLSIKNSNPFVIN
KTKGNSITLDNISVTNNPISNDPKQFIVINDLSSEIINNTGQFNNINPKSL
IIONVPSDVKILNQLVNLINLTAVNSVLSQINNINVOAYSPSINTNS
NINLNSDSSKILPNVILVINSKQLOKITHLANNLNFQDNLIPDPSNSIFPK
NMLTWTVTKNSILKEANFSAAILTKTDPNSNILDSTIKSAKIIDQAGNNSDLNAD
PRTATKQSPKAKTSGWKYE"
complement(3910..3973)
/note="RBP02, RS3-like repeat"
complement(4050..5009)
/gene="RC0851"
complement(4050..5009)
/gene="RC0851"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AA103389.1"
/db_xref="GI:15619954"
/translation="MSKKQDDQIIEKEERLAKLQTPPSLEEEERPAKIQDPPYI
PKKSGNEVDLIELOQSELEKSKSLNDQDQALRDLRLKREDSITTKKPID
LSPNDKAITITDIAQADIASVYANRSTEEKKALHLANLEARDIIRIDIPIS
KEAPKLDIVASWIGKTEBGIKIASPVAIVIKETGSLNYSIKSLPQSSKQKIK
KLIVRYVDLKGIKPKIKSEPLKHYERIDQIPREIIVETAKITQSIIRINGKQKIN
TKQQQIINRMQSPAFINNSSTPSTPNKAKSSKNOR"
complement(5047..5136)
/note="RBP06, A and T rich repeated element"
5281..5332
/note="RBP03, repeated element"
complement(5684..7465)
/gene="pnpA"
/note="synonym: RC0852"
complement(5684..7465)
/gene="pnpA"
/codon_start=1
/transl_table=11
/product="penicillin-binding protein"
/protein_id="AA103390.1"
/db_xref="GI:15619955"
/translation="MNLKLIHGEIISRPAFIGCKGFLSLGIRNPFYLLISBE
YTLSDKRNINFPVLPTRGRIVLDGNTLANKCYQVIVDRSINNRRDLELTISN
IINTSEKYNVYIKOKIKSSRTPPLIIPQDLMQVSNIEBKHLASIFIDIGYLF
YFSSVYSHLIGVQINHEKQELNIIISDFNIGSGIEKYVNNKLRGAFGYKVE
VNAVGYQVREIAGTPTKSGEDMHLNIDVSLQONIQOYINPKSSAIVNDTKGVLIC
ASTPGESNHFESKISENWMQSLTSDPYRLINKVIOYNSPPSVKRIITVLAALVGI
NENKTVPCDGSALGTNSRCNHSQGHGLDMGSLKSCNTYWEILARIYGPDKILE
VAREFVGSKTIDILAPSSGVPSPKEMKKKYLKPSIGDSFNALIGCGFVGTVMQ
LARFTALASNGKLYPRLIKNDPEFVNNIKPNTIIGSLVNTVAVAGTAYNR
ILGNRQLAGKGTQVOGKLNADKLRKSIAMRRHNAFLGAPYHDBSVYTF
IDHGGGSKAAPVARKIMSDVLDKYL"

```
gene complement(7455..7976)
CDS /gene="RC0853"
    complement(7455..7976)
    /gene="RC0853"
    /codon_start=1
    /transl_table=11
    /product="unknown"
    /protein_id="AAL03391.1"
    /db_xref="GI:15619856"
    /translation="MKIKQEIYKNSLOIILKLPALICPTILFPLVSKINQSR
    IFAHEIIFLYFMSLNTSIFPLGFIIDQISGPIGMSLVFSLAVTIYKLSKY
    FAKNYLINFIIICVYCLFIIINFKYLIITIKLEADGYLIIFQFLTIIFSYNIRLV
    LSPMYFKYKAK"
    8041..8068
    /note="REP03, repeated element"
    complement(8226..8420)
    /gene="RC0854"
    complement(8226..8420)
    /gene="RC0854"
    /codon_start=1
    /transl_table=11
    /product="unknown"
    /protein_id="AAL03392.1"
    /db_xref="GI:15619857"
    /translation="MSKNKITDKPSIFIPWIKSPDLTTTINNTISKQLFISSEAO
    EBLEISDNDKRNPSCELL"
    complement(8595..10280)
    /gene="pbbp2"
    complement(8595..10280)
    /note="pbbp2"
    complement(8595..10280)
    /gene="pbbp2"
    /codon_start=1
    /transl_table=11
    /product="penicillin-binding protein"
    /protein_id="AAL03393.1"
    /db_xref="GI:15619958"
    /translation="MKQIMKWKPKIKSLITIMVCSKTKRILMTVCPSFLFECTIS
    VLLIATSVYKKNINSFKKQKPKKEIVDRNGMLIANLSASLFPANPOVIDPETS
    VKKAEILPDIKAKLAELKNSKSFIVKADVLPDQEKITTSIGLGFEEBQKI
    YTESNLHSHVGVGRDVGSLGELAYDKYLTNSDYELNPKPKKEPLQSIDIRQ
    SLSEELDKTKQKFAIGAVGIADPNNGELIALVKNKPEPLAKPEELNINS
    LGIYMGSVFKALIMAVGFDIGVINMDAYDISVWVGFGDLKDYTRQGHVSEIF
    LYSNIGTSOIMLEIGKSNPKYKLGGLDLOLELBERGTPLPSPSKRNELSVT
    MSYGISISPLHFYAMLPAVNGSTLYDLILIKRQEKVIGTVPSEBNTSLOKXIF
    RAVVEGNGKRAEYKYLIGKTKTAETLSOGAGGKKYLNKSSASSTLGLPKSNQ
    YIIFRDEPKTKESFGRATSWTAFTAGVFERMISLYGLEPIEGBEBS"
    complement(10325..10355)
    /note="REP02, RS3-like repeat"
    complement(10325..10355)
    /note="REP02, RS3-like repeat"
    10418..10443
    /note="REP02, RS3-like repeat"
    complement(10447..10472)
    /note="REP02, RS3-like repeat"
    10511..10634
    /note="REP04, repeated element"
    complement(10624..11022)
    /gene="RC0856"
    complement(10624..11022)
    /gene="RC0856"
    /codon_start=1
    /transl_table=11
    /product="unknown"
    /protein_id="AAL03394.1"
    /db_xref="GI:15619959"
    /translation="MSIRKLAHYLLIILIIICSLFSIKERVSTLDYQLSVVKQINS
    ENNNIHLKAEKAYLLPARLEKLAAYLKLKLETVKSYQMIMDPLGNIDQIKNNHI
    SISKSKRYKTIYNKCIQTVSSRVKQTR"
    complement(11029..11952)
gene
Query Match 7.1%; Score 338.6; DB 1; Length 12477;
Best Local Similarity 55.7%; Pred. No. 4.5e-34;
Matches 793; Conservative 0; Mismatches 589; Indels 41; Gaps 6;
QY 2437 TTCATACATAAATAATATACATGATGTTATTATCAGAAACATACCAATCCAGAT 2496
```

```
Db 2079 TGCCTCTCATATTCCTTCGTCATTTATCTATTGTAGAGGATATGTAAAGCTTATATA 2020
QY 2497 TTTCATTTACGGGTGATGCTGAAGCTTCATTCACAAATTCAGAGTTCGTGTAC 2556
Db 2019 TTTCATTTACGGGTGATGCTGAAGCTTCATTCACAAATTCAGAGTTCGTGTAC 1960
QY 2557 ATTTACCTCCAAACGAGATACGAGTTTAAATACCTGCTTTCAATTTCTTACAGAGC 2616
Db 1959 ATTTATTTTAAACGAGCTCTTTGTGCGGATGATATGACGCTCAAAAACCTCTGAC 1900
QY 2617 TTACGATAGTGAACGAATACTTAGACCTTAAACATCACTT-----AATAGGCA 2670
Db 1899 TTCTCAAGGCGATGATATACATTCACAAAGCTATCTCAAGTTTGTATAGATTACACT 1840
QY 2671 ATTAGATATATATCTACTGCTTAAATCTTTGCTGTATCATCTCTTACGA----- 2722
Db 1839 ATTTGCTAATATTTTCAGAAAGTATCATATAGGTTTATATTTTATTTACCTGTAGT 1780
QY 2723 -----GATTTATATTAATAATTCATTCATTAACATTAACATTAAGTTTACCTG 2778
Db 1779 TAAGGATTCATATTAAGTCCATTCACAAAGCTTACGTAATTTGCTAATATTCGCCG 1720
QY 2779 CATTATATGACTGCATTTGCTATTAAGCTTCATCACTTAATATACATGAGTACAAAC 2838
Db 1719 CCATTCATAGATTTGAAGAGCAGAGATGTTTCATGCGCAAAAGCTACGTTCTTTAAAC 1660
QY 2839 TATTTTTCAGATGCTATATCAATAAATACCTACACAAATTCGATATCTGTACAAAT 2898
Db 1659 TGAATATTTGAAGTTCCTTACCAAAATTCCTTAAGAGTGGGATATCTCTTTCT 1600
QY 2899 TTCTACTAAAGAGGATCTCTAATTTGAAGAGATTTAATCTAATATTAATCTCTACA 2958
Db 1599 TTCTATTAATGAAGAGTATCTTTAGAGAGTATCATTAAGCGATATATAGATCTTCTAG 1540
QY 2959 AATCCACCAAGCTTTACTTCACTTCATTCATTCCTTTGAAGAGACAAATTAATCTCAG 3018
Db 1539 GAATTTGCCATTTATACCTTCATCTTGATTTTGAAGATTCGGGTATATTTTAT 1480
QY 3019 ATCTATACATACAGAGATCTTACTATTTTCCCTATATTTTCCCTCGTAATATCT 3078
Db 1479 ATCAATTTTATATTTTTCGCCAAGTTTGA---ATAGTTTATCTTTAAGAAATTT 1423
QY 3079 GATTAATCTTAATTTGGGATATCGATCTGAATCTGTATCTTATCATATTAATACCTAC 3138
Db 1422 TAATTAATTTTACTTGATAGGAATAGGAATATTACGACCTCATATATTAATAGTACC 1363
QY 3139 ATGATTTGCTGCTCTATATATTCATATAGAGAGGTACTATAGACAAATATATTAT 3198
Db 1362 GTATTGGCAATTTCA-----AGATATAGAGGCTTATATTTGGCT 1319
QY 3199 ACTTTCTCTACACCAATATATTAACAGATATTTAGCTGAGATAGATAGTA 3258
Db 1318 GCTTTTCAATTCGCCGAATATTTCTTGATGATTTTCCGATGACATACAGGTAGGCG 1259
QY 3259 CATAGAT-ATAATGAGATATACATCCCTGGATTTTATGATTTAGCTTACCACTA 3317
Db 1258 TGAATATATGAAGGATTTATTAACCTTTTGAATGTTTATGATTAATCTTGGCGAA 1199
QY 3318 CTTCCTTTGCAACCTGGCGAAGCATATAGATATAGATCTCGATGATGATGCTT 3377
Db 1198 GTTCCTTTACCGCATACGACTTTACCGTGATATATAGACAGCTAGAGGATGCGCT 1139
QY 3378 TATTAATCATCTTCCAAATTAAGTATACAGAGGAGTTCAGCTATATTA---TAAT 3434
Db 1138 TTGCTATTCGCAATTTATTTTAAAGTACGGAACATTCGCTATATTCAGTTTAT 1079
QY 3435 CCTCAATGCTGATTTCAATCATATCTTCTTACATGATATACAGACTATATGCTC 3494
Db 1078 CTATAACCTTGAATTTTAAATATATATTTCCGTTTAACTTTGTACTTACAAAGCTC 1019
QY 3495 TCTTTCACTACTTATCTCTCTGTAAGAGGCTTCTATATATCATTAAGACAC 3554
```

Db 1018 TTGTAAGTAAATACTAATTAATCATATTAACGGTTTTCATATATATCGTAAGAC 959

QY 3555 CCATATGACGAGACTTATAGACAGCGCAATTAATGACCCACTAATCATATAACAG 3614

Db 958 CCAATTTTATAGATTAAGTCTGCTTCTATAGTACCGGACCGCTAATATATATACCG 899

QY 3615 GCAATTAAGAGATACCTTTCTTAAGCTTTTCCAGTACACTTAATCATCAATATGAGATC 3674

Db 898 GCATTAAGAGATACCGTTTATATATCTTAATATCCAAATCCGCAATTTGCTTC 839

QY 3675 CTCTTAACCATATATGCAATATACAACTCAGGCTTTTTCATAGCCATCTTGATCG 3734

Db 838 CTGGAAGCCAAATATCAAGTACAACTGACGAGACCGGTTTTCGGAAGATTTTAAAGG 779

QY 3735 CGGATTAACCATCACTGCTTAATTTAGTACATAATTAATCATCACTTAATATCTTTTA 3794

Db 778 CTGAGTACTATAGAGGACACCTTAGGTTAAACCTTCATCTTCAATTTGACAGCA 719

QY 3795 TTGATTTCTGATTAATCACTTATCATCATCAACACTAATCTTC 3837

Db 718 TGAATTTCTGATTAATCACTTATCATCATCACTAATTAACATC 676

RESULT 9
AF176314 23906 bp DNA linear BCT 07-SEP-1999

LOCUS Zymomonas mobilis fosmid clone 42B11, complete sequence.
DEFINITION AF176314
ACCESSION AF176314 GI:5834286
VERSION
KEYWORDS
SOURCE
ORGANISM
Zymomonas mobilis
Zymomonas mobilis
Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
Sphingomonadaceae; Zymomonas.
REFERENCE
1 (bases 1 to 23906)
Lee, H.J. and Kang, H.S.
Sequence analysis of 42B11 fosmid clone of Zymomonas mobilis ZM4
Unpublished
2 (bases 1 to 23906)
Lee, H.J. and Kang, H.S.
Direct Submission
Submitted (07-AUG-1999) Microbiology, Seoul National University,
JOURNAL
Sams6-1 Shillimdong Kwanakgu, Seoul 151-742, Korea
LOCATION/Qualifiers
1. 23906
/organism="Zymomonas mobilis"
/mol_type="genomic DNA"
/strain="ZM4"
/db_xref="taxon:542"
/clone="fosmid 42B11"
104. 4699
/gene="gltB"
104. 4699
/gene="gltB"
/note="glutamate synthase large subunit"
/codon_start=1
/transl_table=11
/product="GltB"
/protein_id="AAD53891.1"
/db_xref="GI:5834287"
/translation="MSIKSSADSPDFVFAERELAKGMYRPEFETDAGCGVLAATDQ
KASRVVTAIDALKAVMHRAVDAGTGGAGAHNDLPASFPDDAIAAHGAHPLN
BLANGVPRPRDLSAQTCTIVSEITIDGRIYGRQYPIVDSVIGEKATRE
IQIYTAGHPDMSANFERKULYIRRIEKKVIAQTNKIISARCLVRSIVYKGL
IAEALPTCYLDQDERFVSRAVFFQRYSTVFPKWMIAQFPEMLANGELINTGR
NMWKSHEIKMASLARGENSEDKPLIPAGASDTALDAVEFTICRAGHAPAAKLMJ
PEAMOHNPMPKARHDNYSYLASVMEPMGPAALITDGSVAAGVDNRNALRPLRTM
TDNLIIYSGESGAVVPENNYIAKQSGPGMTAVNIDSGRFAEDRPLKLIASEP
YGEWIKDEPKLSDLDKSYDPIPYDRDELIKROVSAGLTLEMDIVLSPAEKGA
MSMGDDPTPLVATEKPRITSGPFRONSQVTPNPIIDSLRETQVAKLITRAGNGTL
DSENRKQILITIDPYLKTSEKQKLDHFGEOVAIDCTSPDGAALRLRANISNLT
QENAVRSGHSEFLTDEHISDRAVAVMLATVAVHTHLRKGLRSASTINVASSC
IDPHVAVALIGVATTVAAYLAESAIAVRHSGGLFGTWLTLEQCLDNHHAICNGLKI

MAKGIJAVSSYNGAVNPEAVGLBALVNDLPFGWAKISGEGYASLORSALLRHONA
PAANKLIDVGGYRORASGDSHAYAEIMRLKORXVSQSDSYTYVFERKEHOIDPV
YLDLIGFMSKOPITIEVEPTTEIRRRPTFGNPLGALSPAHETLILANKRITAK
SVSGEGEAPERHPYENDNANSPIKQVASGFRVTAAYLAGCEIEIKVQAGAKG
EGQQLPGFVTDIAKRSTPGVTLISPPPHDLYSIDLAQLIDYDKQINPNNAVQ
VKLSASAGIGITIAAGVAKAHADSIMISPPHVGASPLTISIXAGTPMGMSEANOV
LTLNGRHRVKTARDGLTGRDIVAJAILGAEEFGITSLIAMGCLVROCSYK
PVGICTODRAMQKFGTPERVYNMTFJAEVREILAKGVRSIKEIIGOTDLISQV
NRGADLIDLDLNPULAKVDANTEARVPOIDFERNVPSILDAOMQDAQVFTREK
TOLTVSRNDRALGTRPSAETRNQMSLADGHHVHPTRGSAQSIGAFSVKGLTL
EYFGDANDVYKGLSATTIVRPTVSSPIDGHTNTIIGATVYXGATACGLPAAGRGE
RFAVNRSGAEIVIEBCGAGCEYMTGLVALIGSVGPNAGMTGMAVLDEDSFE
RFAVNPNIIMORLASHEWTVRSMIADAVVITDSFWARVILDEWDMQKFWQICPK
EMLGRLEYPLDDKNTEDQKEKAKGPELKEQEGAF"
4779. 5450
/note="hypothetical protein; ORF1"
/codon_start=1
/evidence="not experimental"
/transl_table=11
/product="unknown"
/protein_id="AAD53892.1"
/db_xref="GI:5834288"
/translation="MMQLYQPLCPFSRKYRFLGKGIYELIRESFPLKDEPYDM
NPGQTPVWVDSAPVILVDSOVCIELEETDRMPLISGTVHRAEIRRLVSLPKN
LYPEYAPLNMERMLKRLVHRQAPDAIIRGAKRANASHLDYIDMLDHRRLAGVY
SLADFAPAAQISVSDYLGVDWRGHNREMYAMKSRPSFOPIIHERMEVILPPDHY
DQDPF"
complement (6178. 6519)
/gene="hima"
complement (6178. 6519)
/gene="hima"
/codon_start=1
/transl_table=11
/product="Hima"
/protein_id="AAD53893.1"
/db_xref="GI:5834289"
/translation="NMRKYDNRNTGQSMQADIPITRADITMLYHEVLSRADSAM
IEMGLHITDALKKENVXISGFSFILDKNERVORBNKTSIEVPIABRVLTFRAS
OLMRORIKGA"
complement (7332. 8759)
/gene="ntrX"
complement (7332. 8759)
/gene="ntrX"
/note="nitrogen assimilation regulatory protein"
/codon_start=1
/transl_table=11
/product="NtrX"
/protein_id="AAD53894.1"
/db_xref="GI:5834290"
/translation="NALIELVVDDEMDIRELVAGVLDGEVATRTAADSATLQALAE
RRPSLVLIDVWLKSGLDLDELLEBEIKRBDPLPVLYSGHNIDNIAVAIAIQGAVD
IEKPFPAERLLILLVERATELERLKRKDIIGTELEITGNSSAINVRAATLAKVA
GTGSRVLITGPGAVGEVAVARLLHMMSPAAAPFIVASAMMEPEVGEELFGEVDSQ
GVYRPLLEKAGGTLFLDEIDAMPPTQAKLITVLTDSGEMRGESEVAVKIDRVYS
ATKAPLARIETIAGRRREDIYRLNVNYPHLPALABRERIPITLDVYAVHATEEVA
TFQVSEBALAQADIDPGRVQKRVNERITILAPGDLISLIDVMLESEVLDGONS
LTADGGSTRSIIIGTPLREARETFERYLREVOIRPSGNISRTASFTGERSALHRLKX
SLGLAEPNANATSNNDNENNEG"
complement (8759. 11065)
/gene="ntrY"
complement (8759. 11065)
/gene="ntrY"
/note="nitrogen regulation protein"
/codon_start=1
/transl_table=11
/product="NtrY"
/protein_id="AAD53895.1"
/db_xref="GI:5834291"
/translation="MMEDSALSTESTETVDEKKRILMPLSKRKYFLYFRPPANAE
LTFAVLAVTAFSRLTYVGRPAKLSLTVYALLVNTVYAMMLMTLARRILAR
AAKSPFGGGRGHRVRLVAFSTVAAPFLLVITFASLFGQVSVSWFSPSARITLANA
DRAVQSVYENDRRIIGDINAMPNDFRALSGAPLEDPMWSFFITQULINRLNEVAL
LQMPEDITAPILMAGLITDRPLENRLPAKMIALLRKQATSTISSGRLTEAAILLD

ASRLTYIVYRQDLPVLLQASAKTALDYNRLVSSHVLIQIRENGALLVVSLLIVA
IAIYAFVADVVRPISHVAVRIRIGSDLGARTHLKHNRDEIGTLGAFANRANK
LEBOITVTLNVDLIDSRATFAVRSAGVLSVSOHTIQLVLSQKLLKXDEK
ILIGEDLNDIAPELDSLAQAGDSIVHILVDEPPTFAVRVADNNVLFEDDITO
QLADGSAAMADVARILAEIKNPLTPIQLAARLORRYGROITSDSEPTSTTV
ROYGDLRWVDESSFAAMPKPYRKEALSDIGRHLPLHFAHQIGVFAQDLP
LWCDOROLRALTNVKNVVEAIEENDRGSGHWMTIKREDSHILVADVGLP
KESRLTEPMTVRGTGTLVAKVIEEHAAGLTFSDRGGGITVLSFDLVALA
PTLISVAEKGIONEPALPGLTRSE"
/gene="nifH3"
/complement(11516..12550)
/codon_start=1
/transl_table=11
/product="NifH3-like protein"
/protein_id="A053896.1"
/db_xref="GI:5834292"
/translation="MNSPEHNPHTNLISPIQISVRIIDNPVILAMTGVTLDPFRIV
RBYGSLANTEMIASQAMIRETROSLQKSTWDPYEPISVQVIGCTPDPMESATLQ
DRGAALIDNMGCPYKTVSGDAGSLKMDLKLASILEATKASLPTVTLKRWGMD
LNSLNPETARINQDVGQVTVHGRTRQILKGAHDAFVQVDAVSIPIVWGD
NSLDAKTALQOQSGADGIVGSGYKPMWLSQMLQGLTDGMEKAPSIDBYRITTE
HYDMLSYGTITGVNARKHIGWYTKGLPASAEFANKINREDEDDRVKRNAGRILPT
LAQPNRLTD"
12702..13865
/note="hypothetical protein; ORF2"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="unknown"
/protein_id="A053897.1"
/db_xref="GI:5834293"
/translation="MMKRTDNLALIVAGOVNVOYKALOKORYKTAGKANTAHADNT
LAPREIDVQVYADGHQTLYOBAYVGRDLPQPVIGVTRRBSVINGKLKAHDKYR
VLHDAARPLKTVLIDRLDLAKSKALPYLPVADPLVNOEVAVRNPLPHRYQYR
QABDELTVAHQAMATGSDPEPDQAVAFKELALVTGDLLEKLTYPDAVSAEA
OMTEKMSVSGGFVDFCEADHMLGKIKIPHDHAGHSDADVALHALTDALGA
IADGIDGTFPSPDPMKGNASTQFLRYVALAKKAGAILDHADTVLCEAPKVPYR
PAWRKILPFEWGLNKSGASKORQLKSLVLAQAKGSLPKRLPVACRIFCAN"
13873..14412
/note="hypothetical protein; ORF3"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="unknown"
/protein_id="A053898.1"
/db_xref="GI:5834294"
/translation="MNPVTRHSIALVACVGTGMAAATAEAGSOSQCTHAEKALM
ETRLPDLQAKQCSVLDSYSLKQKATLSKFRQSESDMKAKAFERLGGSA
VADMGAGRAALITWVSDIVSQKMKPQSCETVNGFLPVAQOPANKIGDVSVTFAL
LKKIGVGRILPFDICSGD"
14327..15103
/note="hypothetical protein; ORF4"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="unknown"
/protein_id="A053899.1"
/db_xref="GI:5834295"
/translation="MFLPSPCLRLKRLADVLCILLIIVREIKPTWMDNNDGDSQ
PSDIEKALQAKQODISADSEKLOKTIOMNEENQOENVHVLPELTPRASLVI
DANRAAGRIIAVASTGGLVMAALTEVPGASDVDAQGVYANAKIDILNIGDVI
ETRGYSILAVMAAAMAYEKSDADIAVLTIGAPTGDEKPKGTIVVPARADAD
PHEVYAEKSPDILGRSGIRLQALCALSLMPDSISQG"
complement(15290..15733)
/note="hypothetical protein; ORF5"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="unknown"
/protein_id="A053900.1"
/db_xref="GI:5834296"

Query Match 4.9%, Score 231; DB 1; Length 23906;
Best Local Similarity 49.9%; Pred. No. 1.3e-20;
Matches 693; Conservative 0; Mismatches 660; Indels 36; Gaps 3;
CDS
gene
/translation="MPHYTESKILPYTPQQLPDLVADISHYPERLPVAVIRISREE
NRMATDVGKAPRESFTKTVLDSHSVSEVITDPLLNHEMFTTEBEKTRL
DFWDFPFRSIFPALGOFEPDAVQKCTQAFERRANNLGLSDA"
complement(15733..16704)
/gene="lipA"
complement(15733..16704)
QY 2489 CCAGAAATTTTCAATTTACGGTGTAAATGCGAAGTTCCATCCAAATTAAGCGCTT 2548
DB 7402 CCAAGGCTTTTAACTTACGATGTAGAGCCGAGCGCTTCAATCAATAAAGAGCTGT 7461
QY 2549 CGTATACATTTACCTCCAAACGAGATACGATTTTAAATCTGCTTTCAATTCCT 2608
DB 7462 CGAAGATATTTGCTGAGAAACGCGCAATTTGCACCCGAGATATTCCTGTAAGGTT 7521
QY 2609 TCACGAGCTTTAGCTAATGTCACAAATTAATCTTACACTTAAACATCTTAATAGGC 2668
DB 7522 TCGGTCCTTCTCTTAATGGGTGCGCATATACAGCGGCTGATTCCTCGCGCTC 7581
QY 2669 GAATTAGTACTATATCTACGTAATCTTTGCTGTAAATCATCTCCTTAGA----- 2722
DB 7582 AAGCTGTTTATGATGCCAGCAGCTCGAAGGAGCATATCGACATATCCGCTTAA 7641
QY 2723 -----GATTCATATTAATTAATTCATTTCTATAACATTAAGTAATGCTTAAGTTA 2773
DB 7642 CGATGCCCGGTGCAAAATTAATGCTGCTTCAACCATTAAGAGCTGTCAGCTTA 7701
QY 2774 CCGGCAATTCATATGATGATGCTATTAAGCTTCATCACTTAATACATGAGTACAC 2833
DB 7702 CCGGCAATTCATATGATGCTATTAAGCTTCATCACTTAATACATGAGTACAC 2833
QY 2834 AAACCTATTTTATACAGATCTATTAATTAATTAATTAATTAATTAATTAATTAATTA 2893
DB 7762 ACAGAGCTTCTGTTGATATGAGGACCAAAATTAATTAATTAATTAATTAATTAATTA 2893
QY 2894 CAATATTCATTAAGAGGATCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2953
DB 7822 CGGCAATTCATTAAGAGGATCTATTAATTAATTAATTAATTAATTAATTAATTAATTA 2953
QY 2954 TCACAAACCTTACAGCTTTTACCTTCACTTCAATATCTTTGAGAGACACATTAAT 3013
DB 7882 TCACGAAACGAGCTTACAGCTTTTACCTTCACTTCAATATCTTTGAGAGACACACATTA 3013
QY 3014 CTCAATCTATACGATACAGATCTTATATTTTCCCTATATATTTTCCCTCCTGTAAT 3073
DB 7942 CGGCAATTCATTAAGAGGATCTATTAATTAATTAATTAATTAATTAATTAATTAATTA 3073
QY 3074 AATCTGATATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3133
DB 8002 ACTGTAATATCTTGGCTGTGCGAATGAGGATATCCGATTCATCAAGAAACAGT 8061
QY 3134 GTACATGATTTGCTGCTTATATTTCCATATGAGAGGATCTATGAGACAAATATA 3193
DB 8062 GTTCTCCATGAGCTTTTCAATTAAGCGGAGCGA-----ACCACA 8103
QY 3194 TTATTAATTTCTCACTACCAATATATTAACCAATTAATTAATTAATTAATTAATTAATTA 3253
DB 8104 CTTGATATCTTCTCAACCAATATATTTCTTCCGACCGCTTCTGATTCATCAATCAATG 8163
QY 3254 GAGTACATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3313
DB 8164 GCGCAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3313
QY 3314 ACTACTTCTTTCACACCTGCGAACCAGTATGATGATGATGATGATGATGATGATGATGATG 3373
DB 8224 GCGACCTCTTTCGACCCGAGAGGCTGTTATTAAGAGCGGAGGCTGCGCTGCA 8283
QY 3374 GCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3430

| | | | |
|------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------|----------------------------|
| Db | 8284 | ACCGTTTCAACGAGCCGACGATGATGCACTGGAAATGCGGTTATTTCTTCA | 8343 |
| Qy | 3431 | TAACTCCAAATGCGTATTTCAACTCATCTATTTCTCTACGTAATACCAAGCTCTATA | 3490 |
| Db | 8344 | GTAAGCCGCAATTTGGCGCTTAAAGAGCGTTTCTCGCCGCAATTTCTGCTTGGGT | 8403 |
| Qy | 3491 | GCTCTCTTTACAACTTAATTCTTCTCTGTAAGAAAGCGTTTCTATATATCATATA | 3550 |
| Db | 8404 | GCTCTTTCGACAAACATATGTAACCGTTGCGCTTCAAGAGGTTCTCGATGAATCAACG | 8463 |
| Qy | 3551 | GCACCCATATGACAGACTTATACAGAGTGGCAATATTACATGCCCATATATCATATA | 3610 |
| Db | 8464 | GCACTTGTGCGAATAGCCGTTACGCCGTTCAATTTGCAATGCGGAAATACCAAGC | 8523 |
| Qy | 3611 | ACAGCAAAATAGAGATACCTTTCTTTAAGCTTTTCAGTACACTTAATCATCATATCA | 3670 |
| Db | 8524 | ACCGCAAAAGTGGATACCGCCGTTGATTTCTTCCAAATATTCAGACCGCTCCAGTTTC | 8583 |
| Qy | 3671 | GATCCTCTTAACCATATATCATATATACATACATACATACATACATACATACAT | 3730 |
| Db | 8584 | GAACTTTTACCAAGATCATCATGTAACCAAGGAGCGCTTTCTCTAAAGCTGT | 8643 |
| Qy | 3731 | ATCGCGATTAACCATATCTGTAATTTAGTACATTAATATCATACTTAATATATCT | 3790 |
| Db | 8644 | AAATGTCATCATCTGTCGAGCGGCTTCTTGTGCAATACCTTCAATCATCAAGAACCC | 8703 |
| Qy | 3791 | TTTATGATTTTGATATCATCATATCATATCAACATTAATCTTCAAGAAATATACAT | 3850 |
| Db | 8704 | GCTACCAATTCAGAAATGTCATTTGCTATCATCAACCAAGAAATTTCTAGCCCATATT | 8763 |
| Qy | 3851 | CTTTCCTTG 3859 | |
| Db | 8764 | CGCTCTTG 8772 | |
| RESULT 10 | | | |
| LOCUS | AE011408 | 14211 bp | DNA linear BCT 28-APR-2003 |
| DEFINITION | Leptospira interrogans serovar lai str. 56601 chromosome I, section 217 of 397 of the complete sequence. | | |
| ACCESSION | AE011408 | AE010300 | |
| VERSION | AE011408.1 | GI:24196156 | |
| KEYWORDS | | | |
| SOURCE | Leptospira interrogans serovar lai str. 56601 | | |
| ORGANISM | Bacteria; Spirochaetales; Spirochaetales; Leptospiraceae; Leptospira. | | |
| REFERENCE | 1 (bases 1 to 14211) | | |
| AUTHORS | Ren,S.X., Fu,G., Jiang,X.G., Zeng,R., Miao,Y.G., Xu,H., Zhang,Y.X., Xiong,H., Lu,L.F., Jiang,H.Q., Jia,J., Tu,Y.F., Jiang,J.X., Gu,W.Y., Zhang,Y.Q., Cai,Z., Sheng,H.H., Yin,H.F., Zhang,Y., Zhu,G.F., Wan,M., Huang,H.L., Qian,Z., Wang,S.Y., Ma,W., Yao,Z.J., Shen,Y., Qiang,B.Q., Xia,Q.C., Guo,X.K., Danchn,A., Salt Giron,I., Somerville,R.L., Wen,Y.M., Shi,M.H., Chen,Z., Xu,J.G. and Zhao,G.P. | | |
| TITLE | Unique physiological and pathogenic features of Leptospira interrogans revealed by whole-genome sequencing | | |
| JOURNAL | Nature 422 (6934), 888-893 (2003) | | |
| PUBMED | 12712204 | | |
| REFERENCE | 2 (bases 1 to 14211) | | |
| AUTHORS | Ren,S., Fu,G., Jiang,X., Zeng,R., Xiong,H., Lu,L., Lu,G., Jiang,H., Ding,Y., Jia,J., Tu,Y., Gu,W., Cai,Z., Sheng,H., Yin,H., Zhang,Y., Zhu,G., Wang,S., Shen,Y., Qiang,B., Chen,Z., Wen,Y., Xu,J. and Zhao,G. | | |
| TITLE | Direct Submission | | |
| JOURNAL | Submitted (12-MAR-2002) Chinese National Human Genome Center at Shanghai, 250 Bi Bo Road, Shanghai 201203, China | | |
| COMMENT | Updated information will be available at our World Wide Web site (http://www.chng.sh.cn/lep/). Comments to the authors are appreciated. | | |
| FEATURES | Location/Qualifiers | | |
| source | 1..14211 | | |
| | /organism="Leptospira interrogans serovar lai str. 56601" | | |
| | /mol_type="genomic DNA" | | |

ESEAVSIQNHFPNGINRELKKNVERLCIMTVGSYITANDARDALGKGTANEMVEIG
 DPKRKEBERGYIILKILQTNNGNVTTRTSVLGIERHSLYKXKMSINISDQYTDG"
 gene
 complement (4466..6256)
 /locus_tag="LA2401"
 /complement (4466..6256)
 /locus_tag="LA2401"
 /codon_start=1
 /transl_table=11
 /product="two-component hybrid sensor and regulator"
 /protein_id="AAN49600.1"
 /db_xref="GI:24196161"
 /translation="MWDETRRYLRDLILGLIILSVLAIAIQSPDQKITPDIDISFL
 DLIVVYFEPFLPSLISLYFNKNKLEGRKKSIRYRLTSLFPAVFLPILQPL
 LPSNVYTKLYERFYGIDIEHLSGSETLLINEKPKKRLLEKTRLESFLIQL
 HESLVAGSKNLINDEYVYIYKKEKPIENALYKPLEKNFKIENATSWKEF
 ASYRDPYCVFIRIPPTGNVLOTGIRIHDSERNVSVSTERNVDRADITREKIP
 YVIRLFTSITVAFILAYESLIPARKISRIEILANATOKISMGCDITLRENG
 EIGALIDSQNVKDLKSKDRLNOSQRIAMKEVQGMHEIKNPLPIQLSERIK
 RKNSENKQFEIYVSGDITIKQVRVLEHVFESDPAMPAKILNQLLEPIL
 VVLEHETPKIKATINISKGFQFLDQKIIISVFTMLKNSIAIERKEKEIPEY
 EGSIRISAVLSRKIKRKAIVISIDNGIGISPELKQKFEFYSTKNNNTSGIGLAIY
 QKSVIDHSGHISVDSSMGCSFQIELPVS"
 complement (6270..6533)
 /gene="pfsH"
 /locus_tag="LA2402"
 /complement (6270..6533)
 /gene="pfsH"
 /locus_tag="LA2402"
 /note="Histidine-containing protein"
 /codon_start=1
 /transl_table=11
 /product="phosphocarrier protein HPr"
 /protein_id="AAN49601.1"
 /db_xref="GI:24196162"
 /translation="MKKILKINENGTMHARPAVFNCAKSPCEITVKKDVVN
 GKSIGMLMALAPENFKIQVEGKEDEALEALNIVNDFV"
 complement (6505..7470)
 /gene="hprK"
 /locus_tag="LA2403"
 /complement (6505..7470)
 /gene="hprK"
 /locus_tag="LA2403"
 /EC_number="3.1.3.-"
 /EC_number="2.7.1.-"
 /codon_start=1
 /transl_table=11
 /product="HPr (ser) kinase/phosphatase"
 /protein_id="AAN49602.1"
 /db_xref="GI:24196163"
 /translation="WSPGINVSNLNEHESLGLLAGQKLTNRINMEINRPGLS
 LTGFESPAHRIQ:FGKEWAVYTSRPEDELEKIAEFEGFHLNCITFGNMPPI
 FVNECKELGIPLMISEVSTKFTLISGLRSLAPTRMGVLIYVGGIILISGS
 GVKSETALIELERGRHLVADWVEIRPLSSYLIGCSLLRHHMIRGIGTINID
 IFGISVARDKILIELIHLEMTBGKDRGTGLENPTBELGVGIPILRPVGRGNT
 FIVETAAVNDLRLGNAAQEFNOKLSYLLQGKVERNPTQO"
 complement (7475..8896)
 /gene="rpoN"
 /locus_tag="LA2404"
 /complement (7475..8896)
 /gene="rpoN"
 /locus_tag="LA2404"
 /codon_start=1
 /transl_table=11
 /product="transcription initiation factor sigma 54"
 /protein_id="AAN49603.1"
 /db_xref="GI:24196164"
 /translation="WNLSSLVOKOTOKLVMTODLRGIELPLSTLELSPRINSELV
 ENPMLSEYASERNRTPDLYSRDLKREKNDPLKNSVMSQDPSLDRKAGGTDGS
 DRNKYTESPEKSLSHLMQRLSLKXDEISIGIILSHMNDGFRPIPIPDIC
 TENKINERKVRVLDQIRHLDPIGAGDVETLILQKILKPEDLVHLIRHICID
 LEKLDYKSIKMEIQLSESVSLASEIKQLEPYPATYTPKRPDYVLPDVIAREVDE
 EDIYNDWIMPLKLVNKEYKNILKNAKSDXEYIATKLSMEMVIRSVNORPQLFLFY
 TSAIEMQIAEFKKGIQIRPLTLKDIKLEKMESESTISRTSKYIQTSGIMELKM

FSSGVRSTREGGIESKTIHDLIRNLVKEQSDNPLSDQEIYDAIGKQIEIARTVA
 KTKIKITIPSSGRKRVSLER"
 gene
 complement (8893..9618)
 /locus_tag="LA2405"
 /complement (8893..9618)
 /locus_tag="LA2405"
 /codon_start=1
 /transl_table=11
 /product="ABC transport system, ATPase component"
 /protein_id="AAN49604.1"
 /db_xref="GI:24196165"
 /translation="MSKTFRMDNLKVNRKRVDSFNIRKKEVGLGNGAGKT
 TSEYVSGVRPDSQKVFIDGQVDVDSPMHIAIRGLVGLAQSASIFRLITVAENLEA
 ILENNLSRTEILIKRDDLILQIRVRVNOGKYLTSGERRCEIARLALNPDITL
 LDPEFAGVDPIAVKIDQIVTIGSLKERRGILITLDHVRVRLTKITDPAVIMYGRILIS
 GSTHIDVNDPERRIYILGSDPTL"
 complement (9631..11001)
 /locus_tag="LA2406"
 Query Match 4.6%; Score 219.6; DB 1; Length 14211;
 Best Local Similarity 50.2%; Pred. No. 4.4e-19;
 Matches 676; Conservative 0; Mismatches 634; Indels 36; Gaps 4;
 gene
 2491 AAGAAATTTCAATTTACGGTGTATCTGATACGTTCCATTCACAAATTCAGCACTTG 2550
 3145 AAGCGACTTCATTTCTGTATAGAGAACGTTCAATTCCTAAACGCGGAAGTTCT 3204
 2551 TGATACATTACCTCAAAAGAGATACTAGATTTTAAATACGTCCTTCAAAATCTTC 2610
 3205 GGTACAGTTCCCTCATTTTGTGTATAGTTTAAATATACGTGCTTCAAAATCTTC 3264
 2611 ACGAGCTTTAGTATAGTACAGAAATACTTAGCACTTAAACATCTAATAGCGCA 2670
 3265 TTAGCTTTTCTAAATCTCTTAACCTCAACCACTTGCTTGA-----GT 3309
 2671 ATTAGATACATATCTACTAGTAAATCTTTGGTGTATCATCTGCTAGAGATTTCA 2730
 3310 TTTAAACCTTTAGAGCGCTCTCTGCGCTTGAGAGATACAGAACCTAGCGTCA 3369
 2731 AATTAATTCATCTTAATACATTAACATTAAGTATGCTTAAAGTTACCTGCCATTCATGA 2790
 3370 AATACAAAGCTTTCCATTAACATTTTGTGTTCCCAAGTTACCGGCAAAAGATT 3429
 2791 CTGATCTGATTAAGCTTCACTCACTTAATCATAGTACCAAACTATTTTCTTACA 2850
 3430 TTGTATAGATGAACCGCTTCAATCTATTTTGTAGAAAGTAATTTTCTTCTCA 3489
 2851 GATGCTATTCAATAAATACCTTCAACAAATTCGGATATCTGTACATATTTCTAATA 2910
 3490 GGTTTGGAATATATATATCAACAAGCATGGAATATGGAAGTTCTTCTCTAAGAG 3549
 2911 CGGTACTCAATTTGAAGAGCAATTAATCTAATTAATCTCAACAACCTACACAG 2970
 3550 TGAATTTGTTATGAAATCAATCAATCAATTAATTAATAGCTTCTCTAACTTCCGTC 3609
 2971 TTTTACTCACTTCAATATCTTTGAGAAAGACAAATATTTCAATCTTAATAAC 3030
 3610 TCGAATCGCCTCTTCACTGAAATATTTGACGCGGAATATTTCAACCTTACGTGAT 3669
 3031 AGGAATCTTACTATTTCCCTATATATTTTCCCTCCCTGTAATCTAGTATCTTAA 3090
 3670 CGTCTCTGACTTCCCAATTTTCAATCTTGTCTCTTGAAGATTCGTAATCTTCCG 3729
 3091 TTGCGTATCGTATCGTAAATCTGTATCTTCAATCTAATAATAAGTATGCTTGC 3150
 3730 CTGAGTGAAGAGATCATCAAAATTTGTCGCAAGAAACAAGTCCCTCGGTGCGCG 3789
 3151 CTTATATATCAATATGAGAGATCTATAGAGCAATATATATATATCTTCCCTCACT 3210
 3790 TTGGAATTTTCCGATTTCTGAATTCGTCGCGCGCG-----TAAAGCTCTTTTGTAA 3843
 3211 ACCAAATATATTAACCAAGTATTTTACCTGATGATGATGATGATGATGATGATGAT 3270

[illegible]

```

gene      /transl_table=11  

          complement(6489. .6955)  

          /gene="RP741"  

          /pseudo  

CDS      complement(6489. .6955)  

          /gene="RP741"  

          /pseudo  

gene      /codon_start=1  

          /transl_table=11  

          7043. .7835  

          /gene="sdhB"  

          /pseudo  

CDS      7043. .7835  

          /gene="sdhB"  

          /pseudo  

          /codon_start=1  

          /transl_table=11  

          /product="succinate dehydrogenase iron-sulfur subunit"  

          complement(8318. .8731)  

          /pseudo  

CDS      /codon_start=1  

          /transl_table=11  

          /product="putative acetyltransferase"  

          complement(9071. .9469)  

          /gene="pgpA"  

          /pseudo  

gene      /pseudo  

          complement(9071. .9469)  

          /gene="pgpA"  

          /pseudo  

          /codon_start=1  

CDS

```


| | | | | |
|----------------------------|--------|--------------------|------------|----------------|
| Query Match | 4.4%; | Score 207.8; | DB 6; | Length 349980; |
| Best Local Similarity | 45.4%; | Pred. No. 4.3e-18; | | |
| Matches 1233; Conservative | 0; | Mismatches 1447; | Indels 37; | Gaps 12; |

| | | | |
|----|--------|-----------------------------------------------------------------|------|
| QY | 1030 | ACGCTCGGCGCAAGTTGCTTGGTCTAAAGAAATTGGACGAACATCTTCCGTGTGTAAGACT | 1089 |
| Db | 214691 | AATATACATATATAATATATATATAATATATATATATATATATATATATATATATATATA | 2146 |
| QY | 1090 | TGCTGGTGAATATAAGAACTTGCTGGTGAATCATTTGATATGAAACAATAAATGATGTTCT | 1149 |
| Db | 214631 | TATATAATATATATATATATAATATATATATATATATAATATATATATATATATATATA | 2145 |
| QY | 1150 | AAAAATA---ATATACCATCAAGATATTATAGGTGTCACATTTGGAGCTGCATTAAGAA | 1206 |
| Db | 214571 | AAATATTTTACATATAATATATATATAATATATATATATATATATATATATATATAATATAT | 2145 |
| QY | 1207 | CATATTTGCAATTCATGTGGAATTAATCGGTGAAAAAATTTAGTATATATCGCTGTTC | 1266 |
| Db | 214511 | AAAAAATATATATATAATATATATATATAATATATATAATATATATATAATATATATA | 2144 |
| QY | 1267 | TACGTATTATACATAAGCATGGAATGGAATTTAAACACTATATATACAAAAATCATTC | 1326 |
| Db | 214451 | AAATATATATATATAATATATATATAATATATATATAATATATATATATAATATATATA | 2143 |
| QY | 1327 | AATGATCT--TCATACATTAATTTGGTCCATCATGCTTTGGAGATCTAATATTAACATGTA | 1385 |
| Db | 214391 | AATATATATATAAATATATATATATATAAATATATATAATATAATATATATATAATATA | 2143 |
| QY | 1386 | CAACGAACATTCACGCAATATGGCTTTGGCATAGAAATAGGAAAAAGTATGAAATATA | 1445 |
| Db | 214331 | TAA---AATATATATATAAATATATATATAAATATATAAATATATATATAAATATATATA | 2142 |
| QY | 1446 | ATACATTAATAATNGATCAACACCTAAAGCTTTGAGAGAACAGACTGTAACCACTGA | 1505 |
| Db | 214274 | TAAATATATATAAATATATATATAAATATATATAAATATATAAATATATATATAAATATA | 2142 |
| QY | 1506 | TATCATTAGCAAAAAAATCTTAATGTGCACTACCAATTTGCATATCTATTTACATTTAT | 1565 |
| Db | 214214 | TATATATAATATATAATATATATATAAATATATAAATATATAAATATATAAATATATATA | 2141 |
| QY | 1566 | TACATGGAATATATCTAGATTAAGCCATATCAACATATTATCTTAGCTATCATCC | 1625 |
| Db | 214154 | AATATATATATAAATATATATATAAATATATATATAAATATATATATAAATATATATAA | 2140 |
| QY | 1626 | TTGATTAATTTCAACACATGATTAATAAACGGTCAAAAAATATGAAATACACGTTCAAC | 1685 |
| Db | 214094 | ATATATATATAAATAATATATATAAATATATATATAAATATATAAATAATATATATAA | 2140 |
| QY | 1686 | ACATATATAAAGAGCATCAAAATCCATGCTTAACATCCAACTAAGTATGTAACATTT | 1745 |
| Db | 214034 | TATATATATATAAATATATATATAAATATATATATAAATATATATAAATATATATATAA | 2139 |
| QY | 1746 | ACATTAACAAACTTAAGTATACTTAAATTTCTAGCTAATTATATATAACAAATTCCTAT | 1805 |
| Db | 213974 | ATATATATATAAATAATATATATAAATATATATATAAATATATATAAATAATATATAA | 2139 |
| QY | 1806 | CTCCTATCTCAATTAACCAATTTGTTATCAATATATAAATAGTCAAAATTTAACTTTTACT | 1865 |
| Db | 213914 | TATATATATAATATATATAAATATATATATAAATATATATATAAATATATATAAATATAT | 2138 |
| QY | 1866 | TATTTATATAATATATATGTTTGACACTTTTTCATACATCATGTATATAATGACATPA | 1925 |
| Db | 213854 | AAATATATATAAATAATATAATATAAATAATATATAAATATATAAATAATATATATAA--- | 2137 |
| QY | 1926 | TTTTTACATAAATACTAATAAATAAAGCAACAATACATATTTAAATACATTAACATAACCTTA | 1985 |
| Db | 213797 | ---TATATATAAATAATATATAAATATATAAATAATATATAAATAATATATATAATATA | 2137 |
| QY | 1986 | TTTACTACGTATAATTACATAGATGTCCTGACATTAATAAATTCATATCATCTTGACATAT | 2045 |
| Db | 213740 | TAAATATATATATAATATATATATATATATATATATAATATATATATATATATATAA | 2136 |
| QY | 2046 | TTCAATAATAAATATGATATAACGTGTTTTTTTATGCCCATTAGTCTTAATAAATGCT | 2105 |
| Db | 213686 | AATATATATATAAATATATAAATAATATATATAAATATATATAAATATATATATAAATATAT | 2136 |

QY 2106 AATTACCAAAAAAGTTAAATATACATCTCTGCAACAATATGAGAGATACATACAC 2165
 Db 213626 AAAATATATATATATATATATATATATATATATATATATATATATATATATAT 213567
 QY 2166 GTCAAAATATATATATATATATATATATATATATATATATATATATATATATAT 2225
 Db 213566 ATAAATATATATATATATATATATATATATATATATATATATATATATATAT 213507
 QY 2226 TATATACATCTAGTATGATATATATATATATATATATATATATATATATATATAT 2285
 Db 213506 TATATATATATATATATATATATATATATATATATATATATATATATATAT 213447
 QY 2286 ATGATATATATATATATATATATATATATATATATATATATATATATATATAT 2345
 Db 213446 ATATATATATATATATATATATATATATATATATATATATATATATATATAT 213387
 QY 2346 TAACAATTCAGAAATATATATATATATATATATATATATATATATATATATATAT 2405
 Db 213386 TATATATATATATATATATATATATATATATATATATATATATATATATATAT 213327
 QY 2406 TCATATCTTTTAAAGATGACTTCACATTCATATATATATATATATATATATATAT 2464
 Db 213326 ATAAATATATATATATATATATATATATATATATATATATATATATATATAT 213267
 QY 2465 GTTATTCAGAAATATATATATATATATATATATATATATATATATATATATATAT 2524
 Db 213266 ATAAATATATATATATATATATATATATATATATATATATATATATATATAT 213207
 QY 2525 TCCAT-TCACAAATTCAGAGCTTCGTATATATATATATATATATATATATATAT 2583
 Db 213206 CATATATATATATATATATATATATATATATATATATATATATATATATATAT 213147
 QY 2584 TTTTAAATCTGCTTTCAAAATCTTCAGAGCTTACGTAATGTCAGAAATATATAT 2643
 Db 213146 TATATATATATATATATATATATATATATATATATATATATATATATATATAT 213087
 QY 2644 AGCACTTAAATATATATATATATATATATATATATATATATATATATATATATAT 2703
 Db 213086 AATATATATATATATATATATATATATATATATATATATATATATATATATAT 213031
 QY 2704 TGTATATATATATATATATATATATATATATATATATATATATATATATATAT 2763
 Db 213030 ATAAATATATATATATATATATATATATATATATATATATATATATATATAT 212971
 QY 2764 TCTTAATATATATATATATATATATATATATATATATATATATATATATATAT 2823
 Db 212970 ACTTATATATATATATATATATATATATATATATATATATATATATATATAT 212911
 QY 2824 ATGATATATATATATATATATATATATATATATATATATATATATATATATAT 2840
 Db 212910 AAAACAAAAAAACCAA 212894

RESULT 13
 AB063522/c 349174 bp DNA linear ECT 22-NOV-2002
 LOCUS AB063522
 DEFINITION *Wigglesworthia brevipalpis* DNA, complete genome, section 2/2.
 ACCESSION AB063522 BA000021
 VERSION AB063522.2 GI:25166251
 KEYWORDS
 SOURCE *Wigglesworthia glosiniidia* endosymbiont of *Glossina brevipalpis*
 ORGANISM *Wigglesworthia glosiniidia* endosymbiont of *Glossina brevipalpis*
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 Enterobacteriaceae; *Wigglesworthia*.
 REFERENCE
 AUTHORS Akman, L., Yamashita, A., Watanabe, H., Oshima, K., Shiba, T.,
 Hattori, M. and Akao, S.
 TITLE Genome sequence of the endocellular obligate symbiont of tsetse
 flies, *Wigglesworthia glosiniidia*
 JOURNAL Nature Genetics 3 September 2002; doi:10.1038/ng986;
 REMARK http://www.nature.com/cgi-taf/DynaPage.taf?file=/ng/journal/vaop/
 ncurrent/full/ng986.html

REFERENCE 2 (bases 1 to 349174)
 AUTHORS Yamashita, A., Oshima, K., Furuya, K., Yoshino, C., Shiba, T. and
 Hattori, M.
 TITLE Direct Submission
 JOURNAL Submitted (21-JUN-2001) Masahira Hattori, Kitasato University
 School of Science, Kitasato 1-15-1, Sagami-hara, Kanagawa 228-8555,
 Japan (E-mail: hattori@genome.sci.kitasato-u.ac.jp,
 Tel:81-42-778-9926, Fax:81-42-778-9926)
 COMMENT On Nov 21, 2002 this sequence version replaced gi:19071886.
 FEATURES
 source
 1. 349174
 /organism="Wigglesworthia glosiniidia endosymbiont of
 Glossina brevipalpis"
 /mol_type="genomic DNA"
 /specific_host="Glossina brevipalpis"
 /db_xref="taxon:36870"
 138..404
 /gene="rpsT"
 138..404
 /gene="rpsT"
 /note="rpsT"
 /note="Protein synthesis"
 /codon_start=1
 /transl_table=1
 /protein_id="BAC24442.1"
 /db_xref="GI:25166252"
 /translation="MAITKSKKRSIOSEKRRKNSKSKSVIRSFIRKVNLSLKEK
 DTINSEIEMQSLIDRYAKNLHKONASRLKSLKXIFIR"
 complement (424..1614)
 /gene="rhaA"
 complement (424..1614)
 /gene="rhaA"
 /note="Transport and binding proteins"
 /codon_start=1
 /transl_table=1
 /protein_id="BAC24443.1"
 /db_xref="GI:25166253"
 /translation="MKKFFICKKPFKNPTFSGLLILFCLAFISNTNFKTYNI
 INYPLIKINHDKNSLTNIIVNDIMTFEIEIGIEIKHEMLVGLSKNRALPGIA
 AIGMIFPALIYNPTMEDSISGMAITVADIAVAVGLKILGHSIPSLIFILS
 LAIFDDGALLIATFYSNHIIDQYMLISLIVLITLSINYLAIVTCIYIYIGILM
 ESTPLSGIHSTIGVILIGILMPSYSYSSKSMKSLKSLKSLKSLKSLKSLKSLKSLKSLKSLK
 NSGIFNPFNLSILPFGIFPGLVAGKIGVLSYSVKRSLKSLKSLKSLKSLKSLKSLKSLKSLK
 GISPLCGIGFMSILFISNLAFCINIKIYIAKFSILISISVSVIGFLYELYYKI
 KLNK"
 complement (1675..2799)
 /gene="dnaJ"
 complement (1675..2799)
 /gene="dnaJ"
 /note="Protein fate"
 /codon_start=1
 /transl_table=1
 /protein_id="BAC24444.1"
 /db_xref="GI:25166254"
 /translation="MAKSDYDILIGISKASPREIKTAYKRLAVKHPBPNGNLAE
 SKREIKRAYEVLLDEKRAATVQYHTAFDQSSNTERFSNADPIDEDVGDIFG
 NNRKRSNRSGSLQYININLSLEAVAGISKEITIPYLERCNISGKMPGISTAONCS
 TNGGQGIQWRQGFSEVOQCPSCRCRCKIRPCPNCGSGGVESKINSIRIPAGV
 DNDKIDRLSGEGBAGKLGSGSDLVYKISVKEHPFKREBNHLYCEVPINFAMAALGG
 DIWVPLADRGVVKLRIPAEOTGRLFMRGKGVKISGSGDGLLCRVVETPVRLSEK
 QKSLKEIGISFVGRGMYNSPRSKSPFGVKKFFPGLNN"
 complement (2874..4811)
 /gene="dnaK"
 complement (2874..4811)
 /gene="dnaK"
 /note="Protein fate"
 /codon_start=1
 /transl_table=1
 /protein_id="BAC24445.1"
 /db_xref="GI:25166255"
 /translation="KEILKGNITIGDILGTTNSGVALIENGKXVYNSGDRTPSII
 AYENELIIVGQPAKQSVTPNTEFALIKLGRFTDHEVDVIMPIYKIVSEN
 GDWLVWVKQKVAIPQISAEILKMKKTEADYIGKSLTAVIVPVPVFDQKATKD
 AGRIAGIDVRIINEPTAALAYGLDKTKGNRIIAVYDGGTFDIIISIEIDVDVDEK

TFEVLSTNGDTHLGEDEPDSRLINLYVNEFFKEQIDILNDPLAMORLSESEKAKIE
LSVYQTVNLPYITADSSGPHNNIKYTRAKSLVLELYKTLEPVTKSLDQAKK
LIDKIVLVGGOTMPLVOKKSDPFPEKPRKYNPEAVAIGAAGVGLVADK
VLLIDVTPLSIGETMGVMTLLSKNTTTPKSCIPSTAEQNSAVTHLVGEEK
RSIDPKSIGORVLDGAPAMGMPQIEETTPIDADGIIHSAKADNSGREKTTKAS
SGSENEIDMKKESEANABDIFEBELVKNDADHLHSTRQIIEANPLERKT
EIEKINELSLISIKGEDKDIKIQSLIQISGLVPSKNENKINKEDIKIKNKEN
NKIKNDVVDAAFEIEIKDKKN"
gene 5858. 6391
/gene="fabA"
5858. 6391
/gene="fabA"
/note="Fatty acid and phospholipid metabolism"
/codon_start=1
/transl_table=11
/protein_id="BAC24446.1"
/db_xref="GI:25166256"
/translation="MVKKHKFYTKEDLLCSRGSLPGKYGPOLPAPYMLIMDIRVKT
ENGWYNGKFIKAEIDIPNMFSCHEIDPWPVPCGLGDAMQVGLNGFIHGGG
KRALGVRKVFSGQILPTSKIVYIHFRIIRKLFPGWADGEVFCDDKITYTAND
LKVGLFDITSFKDFK"
gene 5859. 6391
/gene="yhgI"
5859. 6391
/gene="yhgI"
/note="Unknown function"
/codon_start=1
/transl_table=11
/protein_id="BAC24447.1"
/db_xref="GI:25166257"
/translation="MFKYIITINFEYFKKKIIPYKMIISKSAKHYSKLSKKK
GNIKLSTENKDKNTLCQIKYFPDKENENYTKNFIPEVYVYKNIPIYIKKID
IESKIKDEILIKIKKEKENKRIISNITKIKNSKIFEGYKVEYELDIENNFVILK
FEGGNGCSMAKVLTKESIEKEIKKNPNINQVIDITDIHSESSP"
gene 5860. 6391
/gene="pro-TGG"
5860. 6391
/gene="pro-TGG"
/note="Cell envelope"
/codon_start=1
/transl_table=11
/protein_id="BAC24448.1"
/db_xref="GI:25166258"
/translation="MKNKNIIPNLPYSGKEIKYIDPAIRNKTSGSGYPSNLCTL
MIENKNTIKVFLTPSCITSLSEMAALLDKINKDEIIMPSTFVSTNPNVGLGAKIV
FVDIRSDMLINENLIESAITRTKRAIVVHVAIGISEMDFIKNISRYKVLWIEDSA
OSMSKYKVLIGSLIGHGCFSEHETKRYTSGSGGAILINDESLIRASIIIEKGN
RSLFKTKTKYVWADIGSSYTIIDIAVYMSQKISEKIOKREMLYGVLSLEN
IKLPIFIPFNCKCNANTFYLIPINDKTKCNLFIEYMKNNKITTLIRIYVLIHSPAG
KRGKHNDINTIKTSNNILVRLPFLFKLNKNDQCIISNINPIDKS"
gene 5861. 6391
/gene="rpo"
5861. 6391
/gene="rpo"
/note="Transcription"
/codon_start=1
/transl_table=11
/protein_id="BAC24449.1"
/db_xref="GI:25166259"
/translation="MNTLKNKILISDITLIGENVGLNLANRKKODITSTKQAHK
IGEDIPDGVLEIIDQDFGFLRSSDSSYLAPDDIYVSPQIRRFNLRITGDTISGRIR

PEKGRYVALLKVDENVYDKPSSRNKILFENLTPLHANSRLKMERNGSTEDLTTR
VLDLAIPIRGGRG.IVAPPKAGKMLLONIAQSIYNHPOCVLLVILIDEPBAYTE
MORIVGEVIASTPDEPSRHYQVSEMTERKRIYEHKQVITILLDSTTRAAVNT
VYPAAGKUTIGVDANALHRPREFGAARNEEGSLITIALALDITSGKNDVYEE
FPGTGMELHLSRKIAEKRVFPAYIDNYSGRKSELTSSEBQOMILIRIHPONE
IDGMFELNKLSEVFKINIEKSIYQEKLVKSIPSPFLPPEKIIDKXIGIVSQKLE
KISEKYVNLKLEVFKNIEKSIYQEKLVKSIPSPFLPPEKIIDKXIGIVSQKLE
ITIDENLRL"
gene 11604. 12788
/gene="pgk"
11604. 12788
/gene="pgk"
/note="Energy metabolism"
/codon_start=1
/transl_table=11
/protein_id="BAC24451.1"
/db_xref="GI:25166261"
/translation="MKIIVYDNLKDKKVLRSOLNVPINQKIMCMRIKSLIETI
NVIINAKCIVVASHLGNPIEKDYDFSLHIVNHKKIKYRILIONLNEFEA
KEKEIIVLEWVFNPKRNDNLSEKYNALCDIIVMSFGSHSESTTGIIKFPAP
ISIGLFLFKIKYKALNSKRPITVIFGSGISTGLGVKLSNSENVLVGGGI
ANTLIPSGHNGKSLHDKNNISYKIKKELVKNIIIPDPVTDQINISYKESIK
EIKNDYVIDYKNSCECFIKIKAKIKFNNQGLIEIOPFQNGTEINSYVHSS
AKSITGGVIVLMMFNFIDRFISTGSGAFSLERKELPVLSLKKFKYINPK
Q"
gene 12811. 13881
12811. 13881
Query Match 4.3% Score 204; DB 1; Length 349174;
Best Local Similarity 47.8%; Pred. No. 1.3e-17;
Matches 828; Conservative 0; Mismatches 880; Indels 26; Gaps 7;
QY 400 ATAGTACTATCTAATTAATAAATCTCAATGCTTTTGCAAGAACTACAAATATA 459
276130 ATTGATTAATATTTTAAAAAATAATATATATATATATATATATATATATAT 276071
DB ATGAGACTTACTATATATATATATATATATATATATATATATATATATATATAT 519
QY 460 ATGAGACTTACTATATATATATATATATATATATATATATATATATATATAT 276011
276070 AATTAATATATATATATATATATATATATATATATATATATATATATATAT 276011
QY 520 TAATATACATCTGACATGATTAACAAATGATTAACAAATTTCTTAACACTAGTAATATAC 579
276010 TATTTAAATTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 275951
DB AATTAATATATATATATATATATATATATATATATATATATATATATATATATAT 639
QY 580 AATTAATATATATATATATATATATATATATATATATATATATATATATATATAT 275951
275950 AATTAATATATATATATATATATATATATATATATATATATATATATATATAT 275951
DB CAGTATTTAGGTGACAGATCAATTTGGACACAGCAATAGCAATTTGACGTGACAGCATG 699
QY 640 CAGTATTTAGGTGACAGATCAATTTGGACACAGCAATAGCAATTTGACGTGACAGCATG 699
275894 TACTGTAAATTTGAGGTGAGGTGATGAGCAACAGCTGTGCTGTTTCCATTTTCAAAAAA 275835
DB TACTGTAAATTTGAGGTGAGGTGATGAGCAACAGCTGTGCTGTTTCCATTTTCAAAAAA 275835
QY 700 TATATAGTAACTATAGGAGGACGATGATCAATGAATTTTACATATTAACACTACCTACG 759
275834 TAGAAAAGTGTTTTATAGGAGGATGAATAAATAATATGCAATCTCTTAAAAAAGATG 275775
DB TAGAAAAGTGTTTTATAGGAGGATGAATAAATAATATGCAATCTCTTAAAAAAGATG 275775
QY 760 AAAAAATTAATAATTTTACCACATATCATCTACACAGCAATATATGCAACAGCAA 819
275774 TTGCAATATAAATTTTACCACAAATTAATTTTCTAATGATTTAAAAATGAAATTTTC 275715
DB TTGCAATATAAATTTTACCACAAATTAATTTTCTAATGATTTAAAAATGAAATTTTC 275715
QY 820 TATAGCGAATATATCTGACAAATATCATGTTATTTTCTTAACTATCTTACGACAA 879
275714 TTTAAAAAAGCTATTAATAATATATGCAATATC---TATGTATATAGTGTTCCTAGATAG 275658
DB TTTAAAAAAGCTATTAATAATATATGCAATATC---TATGTATATAGTGTTCCTAGATAG 275658

| | | | |
|----|--------|----------------------------------------------------------------|--------|
| QY | 880 | ATTACGACCATATGTACACAAATACACAAACAGCATATGTGTAATACTCCAAAT | 939 |
| DB | 275657 | TTTAAACATTTTAAATTAATTC-----AACATCTTAAATTAAGAAATGTTTTT | 275604 |
| QY | 940 | ATTAAATTTAGTAAAGGTATGGAAATTCATCATCTCAATTTCCAGTGAATATGCGA | 999 |
| DB | 275603 | AATATGTGAACCTAAAGGGTTGAACCAAGAAACAGAAACATTAATCAAGAAATGTTTA | 275544 |
| QY | 1000 | AGAAATTTTACATATATATCAATTTT---TATCTCTGTGTCGCAAGTTTGTCAAAGA | 1056 |
| DB | 275543 | CGATATCTTAGGAAGAAACAAAGTTTGCAATTAATGAGCAACATCTTTTGCAAGTA | 275484 |
| QY | 1057 | AATGCAACATCTTCTTGTGATATGATATGATCTGTGCTGATTAATGAAGAACTTGTA | 1116 |
| DB | 275483 | TCATGATATGATTAACCACTCAATGCTACTACTATATATGATTAATTAACCTTGTA | 275424 |
| QY | 1117 | ATCATGTATAGAAACATATATGATATGATCTTCTAAATAATATATCAATCAAGATATAT | 1176 |
| DB | 275423 | ATTTTATAGCAATTAATCAATTAATTAATTTTAAATATATACAGATTTCTGATTTAGT | 275364 |
| QY | 1177 | AGGTGACATGATGAGCTGCATTAAGAACATTAATTCATGATGCAATGGAATATGCG | 1236 |
| DB | 275363 | TGAGAGCAAAATAGGAGAGTATAAATATGATCGGATAGCTTCTGATATGCTGA | 275304 |
| QY | 1237 | TGCAAAATTTAGTAAATATGCTGTTGCTACTGTTATTAATAAGCAATGAATGAAT | 1296 |
| DB | 275303 | TGGAATACGATGGGTCTTACGCTTAAACAGCTTTATTAATCTTACGATTAAGAAAT | 275244 |
| QY | 1297 | TAAACATATATATAGCAAAATATCATTAATGATATCTTCAATTAATTTGTGTCATC | 1356 |
| DB | 275243 | GTATA-----AGTGGGTAAAGTTATGGGGCCCAATGATACATTTTATGGGAATGTC | 275190 |
| QY | 1357 | ATGCTTGGAGATTCATATTAATTAATGATACACAGAAACATCAACCAATATGCTTTGG | 1416 |
| DB | 275189 | TGAGATGAGAGACCTTGTTCTTACATGATGATATGATATGAAGAAACAGAAATTTGG | 275130 |
| QY | 1417 | ACTAGAAATAGGAAAGGTAGAAATATTAATCATTAATAGATCAACAACCTTAAGCTTGT | 1476 |
| DB | 275129 | GATATATAGCTCAAGATATAGCATAGAAATCTAAATCAAAAGTTGATGATAT | 275070 |
| QY | 1477 | TGAAGAACAGCATCTGTAAACCACTGATATCATTTAGCAAAAAAATTAATGTAACCT | 1536 |
| DB | 275069 | AAGAGATATCAATATATTAATTAAGAAATTTTATTTATCATGTTAAACCAAAATTAACAT | 275010 |
| QY | 1537 | ACCAATTTGCAATCATTTTCAATTTATATAGAGAAATATACATAGATTAAGCAAT | 1596 |
| DB | 275009 | GGCTATATCAACAAAGTTTAAATTTTAAATATGATATGATCAACAGTAAATAATCAT | 274950 |
| QY | 1597 | ATCAACATATATCTTATGATCATCTTGTGATTAATTTCAACCAATGATTAATTAAC | 1656 |
| DB | 274949 | TTTAATATTTTAAACAAATAATTAATTAATTAATTAATTAATTTTAAATTTTAAATTA | 274890 |
| QY | 1657 | GGTCAAAAT--ATGAATACAGCTGACCAATATTAATTAAGAGCAATCAATACCAT | 1714 |
| DB | 274889 | AATTAATATGAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA | 274830 |
| QY | 1715 | CGTTAATCACTAATGATGATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT | 1774 |
| DB | 274829 | AAATATGTTTACCTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA | 274770 |
| QY | 1775 | TTCCATGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT | 1834 |
| DB | 274769 | ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT | 274710 |
| QY | 1835 | GATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT | 1894 |
| DB | 274709 | AGCAAAATTTTAAATCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT | 274650 |
| QY | 1895 | TTTCACTCATCATGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT | 1952 |
| DB | 274649 | ATTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT | 274590 |
| QY | 1953 | ACCAATACATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT | 2012 |

-35_signal

```

terminator      1735..1752
protein_bind    1824..1838
                /bound_molecule="NR 1"
gene            1854..2900
                /gene="ntcB"
                /length="2900"
                /name="ntcB"
                /codon_start=1
                /transl_table=11
                /product="NR 11"
                /protein_id="CAA48235.1"
                /db_xref="GI:45901"
                /db_xref="GOA:P28788"
                /db_xref="SWISS-PROT:P28788"
                /translation="METGNLPDNTLLIDSLHSVLVINEPFIICVANHAAQVLAQSR
                RKLPEPTDLSYHSPDAELMDLTLGQSPFDNEVILVHNSQHTMSLSAQISEQ
                HILLELPMSORLSQEQIQQAQIARELVGLAEIKNPLGGLGAQLAKSLP
                DPALTEYQVLIHQADRLRLVDRLLGPQHGKTHGSIHHVERVAQLISLCPENV
                TLKDYDPSLHSHYDPOIEOVILNTIRNALQAVERTGTTILRLTRAFQITLHGER
                HRLVARDIVDTGSGIPPHLQDTLFYPMWSGREGDNGLSIRANLVDAQKILERTS
                MPNTESITVPIK"
                2911..4332
                /gene="ntcC"
                2911..4332
                /gene="ntcC"
                /codon_start=1
                /translation="experimental"
                /evidence="experimental"
                /transl_table=11
                /product="NR 1"
                /protein_id="CAA48236.1"
                /db_xref="GI:45902"
                /db_xref="GOA:P28787"
                /db_xref="SWISS-PROT:P28787"
                /translation="MOKGVVVDSDSSIRWLEPRAITTRBGMCRAPFHANDVLRALN
                SEQDVLISDIRMPDMGSLIKTIKQVPELVITMTASDIDAAVNAQCAKPYLT
                KPFPIDIBETLALIRATTHREKQKPNNAENTLOSVDYDGEAPMGEVRIIGRLSR
                SSISVLNGESGTEKELVAPLHRSFRALAPFLALMAAIPLDLISELFGHEKAP
                TGAQVQGRERQANGSLFLDELIDMPDLIQRLRLVLEGGQFVGGVAPAKVVR
                IIAATHQDLERKVEGDFREDLVHRLNVIRIQPLPDRDEDIPLSLARYLQTALEL
                GVETKSLHSGSLKTMMEVWSGAVROLCRMVLTWTAQOISIMPOLDPEILADEK
                AKINRLTISQMSOHLISLWDEALGEGKENIINDALPQFERTLLSLAVTQHKMDA
                ARLMGWRNTITRLKELGIEDY"
ORIGIN
Query Match      4.0%; Score 188.6; DB 1; Length 4362;
Best Local Similarity 49.6%; Pred. No. 5.4e-15;
Matches 575; Conservative 0; Mismatches 569; Indels 15; Gaps 3;

```

```

2699 TTGGCTGAATCCTCCTAGAGATTTCATATTAATCCATTCATTAACATTACGT 2758
4052 TCTTGTGGCATGATTTCTGGCTGCAGTCATAACCGTCAGCCACACACATTTTCG 3993
2759 AATTGCTTAACTGACGCAATTCATGATGATGATGATGATGATGATGATGATGAT 2818
3992 AGTGTCTTACGTTCCAGACCAACGATTCATGATGATGATGATGATGATGATGATG 3933
2819 AATACATGATACACAAACCTATTTTTCACAGATGCTATTCATTAATACCTACACAA 2878
3932 AGAGATTGGTTTCAACACCCAGCTCTTTGGCGTTTTCAGAAAAGTAGCGCGCTAAA 3873
2879 TCCGCTATATCTGATACATATTTCTACTAAGACGCTACTTAATTTGAGAGACATTAAT 2938
3872 CTAGGAATATCTTCTGTTCTATCCGGAAGAGTGATATTGATGATGATGATGATGAT 3813
2939 CTATATATTAATCTCTACAAACCTACAGCTTTTACTTCACTTCAATATCTTTGAA 2998
3812 CGATGATGAGATCTTCAGAAATCGCTTCATCTCTGTTTCTTAATCCTGATGG 3753
2999 GAAAGACAAATATCTCAGATCTATCTAAGACGATCTTACTATTTCCATATATAT 3058
3752 GTTGTGCAATATCTCGACATCCACTTTAAGAGGCGATGACCTCCAACTCAGATAAT 3693
3059 TTTCCTCTCTGATATATCTGATATCTTAAATGCGATGATGATGATGATGATGATGAT 3118

```

```

Db 3692 TGCCCTTACGCCAATACCGTAAATTAACGAGTTTGATATCATATGATATCAACAATT 3633
Qy 3119 TCATCTATTAATTAACGTCACATGATTTGCTGCTATATATTCGAATATGAGGAGTACT 3178
Db 3632 TCATCCAGAAAAGAAAAGAAACCAACATTTGGCTTTTCAAAACGCC-----TTGGGTTCT 3579
Qy 3179 CTATGACAAATATATATTAATCTTCTCTACTACCAAAATATATTAACCAAGTAATATTA 3238
Db 3578 TGAGAGAGACACAGTAAGACACCTTTTCAATGACCAAAAGATGATGATTAATTAATCT 3519
Qy 3239 GGTGTAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3298
Db 3518 TTGGGATGCTGCTCCATATTTTGGCCATTAAGAGAGCAAAAGACGAGGCTATGACGA 3459
Qy 3299 TGTATTAACCTAGACACATCTTCTTTCCAAACCTGCGAAGCAAGTAATGATATGACGA 3358
Db 3458 TGCAGTGGGTGAGCAACAGCTTTTACCTGTAACCCGATTCCTCATTAATGACACACTA 3399
Qy 3359 CTCGATGATAGTGCCTTTATATATCTACTCTCAATACGATATACAGGAGGAGTGA 3418
Db 3398 ATGATAGAGCGAAGATGACCGGTAATACGATTAATCTTTGATGACAGGAGCTTCG 3339
Qy 3419 CCGACTATTTTCATATC--CTCAATGCTATTTCAATCATGATTTTCTCTACGTAAAT 3475
Db 3338 CCAATCATATGAGAAACAGATTAAGATATTTTCAAGATATTCGGCTGCTTTGGCTCT 3279
Qy 3476 CTACAGACTCTATAGCTCTCTTTTCAACTATTAATCTTCTCTCTGTAAGAGCTTT 3535
Db 3278 CGATAGTGTGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3219
Qy 3536 TGTATATATCATTAAGACCAATATGACAGCTTTTACAGAGTGGCAATATTAACATGC 3595
Db 3218 GGTATGATATCAAAAGACCTTTGATTAAGACATTCATGCGGATCAAGATCAGAATGT 3159
Qy 3596 CCACTAATCATATTAACAGCAAAATAGATACCTTTTAAAGCTTTCCAGTACACT 3655
Db 3158 GAGTCATATTAATTAACGAGGAGTGTAGATATGTTCTTTATATTTTAAGTAGAT 3099
Qy 3656 AATTCATCAATTCAGATCCCTTTTACCATTAATCAATTAATCAATCAATCAATCAAT 3715
Db 3098 AACCATTCATATG-----GCTAGCAATGCTGATATATATATATATATATATATAT 3045
Qy 3716 TCATAGCATCTTGTATGCGGATTAACCATCACTGTAATTTAGTACATTAATATCA 3775
Db 3044 CTATTTAGCGCTTGTAGTACATCTTGTAGCATGTTCAAAAGACGACATATACCTTCT 2985
Qy 3776 TCATTAATATATCTTTATTAAGATTTTGTATCAACTTCATCATCAACAATTAAT 3835
Db 2984 CGTATATTTGGCGCTTCAAGAACCCAGGAAATTAAGCTGTATCATCAACACCATCA 2925
Qy 3836 TCAGAAATATACATCTTT 3854
Db 2924 TTTCCTTTTGTGATGTT 2906

RESULT 15
AX599046/c 8056 bp DNA linear PAT 14-FEB-2003
LOCUS AX599046
DEFINITION Sequence 386 from Patent WO02077272.
ACCESSION AX599046
VERSION AX599046.1 GI:28399186
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1.
AUTHORS Berlin, K., Braun, A., Distler, J., Guelzig, D., Howe, A., Mueller, J.,
Olek, A., Piepenbrock, C., Adorjan, P., Grabs, G., Liesche, R., Leu, E.,
Lewin, A., Lipscher, E., Mater, S., Model, F., Mueller, V., Otto, T.,
Piet, C. and Ziebarth, H.
TITLE Methods and nucleic acids for the analysis of hematopoietic cell
proliferative disorders

```

JOURNAL Patent: WO 02077272-A 386 03-OCT-2002;
Epigenomics AG (DE)
Location/Qualifiers
Source 1. 8056
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="chemically treated genomic DNA (Homo sapiens)"

Query Match 3.98; Score 186.4; DB 6; Length 8056;
Best Local Similarity 43.88; Pred. No. 8.2e-15;
Matches 1900; Conservative 0; Mismatches 2341; Indels 99; Gaps 21;

QY 308 TCTTCCTGTCATCATATTTTCCATTTATTTACAAATATCATAGAGTCTTTATTCG 367
DB 4643 TTTTATTTTAAAAAATTAATTTTAAATTTTAAAAAATTAATTAATTTA--AA 4586
QY 368 ACATCATATTAATATCATCAAAACAAATTAATAGTACTATCTAATTAATTAATC 427
DB 4585 ATTAAATTTTAAAAATATTAATTAATTAATTAATTAATTAATTAATTAATTA 4526
QY 428 CTCATAGTCTTACGAAAGACTCAATATTAATGAACATTAATTTATATATTT 487
DB 4525 AAAACAAATATATCAATTTTATTAATCAATTAATTTAAAAAATTAATTAATTT 4466
QY 488 CAATTTATTAATCTCATCATATTTAGCATTAATATCACTGACATGATTAACAT 547
DB 4465 CATATTTAATTTAATTTTATTAATCAATCAAT--TTCAATATTTTACATTCGAAT 4408
QY 548 AGTAAACATTTTAACTAGATATATATCAATTAATTAATTAATTTTGAATGCT 607
DB 4407 ACAAAAAATTAACAAAAAATTTTAAACATTTTAAATTAATTAATTTTAAAT 4348
QY 608 TTCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 667
DB 4347 AATTAATTTCAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4288
QY 668 ACAGCAATGCAATGCACTGACGACATGATATCACTTAACTTATGAGGAGCTGAT 727
DB 4287 ATCAAAAAAATTAATTTTATTAATTAATTAATTTTATTAATTAATTAATTA 4228
QY 728 CA-----TAGAATATTAACATTAATTAACCTTACGAAAAAATTTAAATTAATTA 783
DB 4227 TAAATTTCAATTAATTTATTTTATTAATTAATTAATTAATTAATTAATTAATTA 4168
QY 784 ATATCATCTACGACACATATATGCAACGCAATATTAATTAATTAATTAATTA 843
DB 4167 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4108
QY 844 CATATCATGATTAATTTTAACTATTTCTCAACACATTAATTAATTAATTAATTA 903
DB 4107 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4048
QY 904 ACAACACAAACGATATGTAATTAATTAATTAATTAATTAATTAATTAATTAAT 963
DB 4047 ATAAACAAAA-----AATTTCAACACAAACAAATTTAATTAATTAATTAAT 4000
QY 964 AATTAATCATCAATTTTCCAGTAATTAATTAATTAATTAATTAATTAATTAAT 1023
DB 3999 TTTAAATTTTATTTTATTTTATTAATTAATTAATTAATTAATTAATTAATTA 3940
QY 1024 TTTTATCTCTGTCAGTATTTGCTAAGAAATTAATTAATTAATTAATTAATTA 1083
DB 3939 AATTAATTTTATTAATTTTATTTTCAATTTTAAAA-----CATTAATTAAT 3887
QY 1084 AGTACTGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1143
DB 3886 ATTTCAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3827
QY 1144 TGTCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1203
DB 3826 AAAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3767

QY 1204 GAACATATTTGCAATTCGATGGAATATTCGTCGAAAAAATTTAGATTAATTAATGCTGT 1263
DB 3766 AAAAAATTAATTAATTAATTAATTAATTAATTTTATTTTATTTTAAATTAATTTAA 3707
QY 1264 TGCTAGTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1323
DB 3706 TCATTTTAAATTAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3647
QY 1324 TTCAATGATCTTCATTAATTAATTTGCTCAATCAATGCTTGAAGATTAATTAATTA 1383
DB 3646 TACATTAATTTTATTAATTAATTAATTTTATTAATTAATTAATTTTATTAATTAAT 3587
QY 1384 TACAACGAAATTCAGCAATATGCTTTTGAATTAATTAATTAATTAATTAATTAAT 1443
DB 3586 TCATTTTATTAATTTTCAACAAATTTTATTAATTAATTTTATTAATTAATTAATTAAT 3527
QY 1444 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1503
DB 3526 AATTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3475
QY 1504 GATATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1563
DB 3474 ATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3415
QY 1564 ATTACATGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1623
DB 3414 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3355
QY 1624 CTTTGAATTTTAAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1683
DB 3354 ATTCATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3295
QY 1684 CCACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1743
DB 3294 ACABAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3235
QY 1744 TTACATTAACAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1803
DB 3234 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3176
QY 1804 ATCTCATATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1863
DB 3175 ATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3116
QY 1864 CTATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1923
DB 3115 ATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3056
QY 1924 AATTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1983
DB 3055 ATTTATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2996
QY 1984 TATTTACTAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2043
DB 2995 AATTTTCAATTAATTTTATTTTATTTTAAATTAATTAATTAATTAATTAATTAAT 2936
QY 2044 ATTTCAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2103
DB 2935 TCATTAATTAATTAATTAATTTTTCATTTTCAAAATTTTATTTTTCACAAATTT 2877
QY 2104 CTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2163
DB 2876 -TTTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2818
QY 2164 ACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2223
DB 2817 ATAAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2758
QY 2224 AGTATATCATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2283
DB 2757 TTTTATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2703

QY 2284 TTAATGATATTTATCAATTCACAAATTTCCAACTTCAACAACAACAACTCAAGTCTAT 2343
DB 2702 TAAATTAATTTTAAATTTTAAATTTTATTTTAA-----AATTTAAAT 2654
QY 2344 TATAACAAATTCAGAAATTATATATCTACCAACACATAGTATCAAAATTCATATA 2403
DB 2653 TTTAACATTTTATTTTATTTTATTTTAAATTTTATTTTAAATTTTATTTTCA 2594
QY 2404 CATCATTCCTTTTAAAAAGTCACTTCAACATTCATACATAATTAATTAACAGCA 2463
DB 2593 TTTATTTTATTTTATTTTAAATTTTATTTAAATTTTAAATTTTATTTTATA 2534
QY 2464 TGTATTTCAGAAACATTCACATTCAGAAATTTCAATTTAGCGTGAACG 2523
DB 2533 AATTATTTATTTTATTTTATTTTCAATTAATTTTAAAT-----AATTTAAACAAAT 2478
QY 2524 TTCCATTCACAAATTCAGAGTTCGTATACATTCCTCCAAAGAGATACAGT 2583
DB 2477 AATTAATAATTTATTAATTAATTAATTAATTTTAAATTTTAAATTTTAAATTT 2418
QY 2584 TTTTAAATTCGTCTTCAAAATTCCTCAAGCTTACGTATGTATGATACAAATTCCTT 2643
DB 2417 TTTTATTTTATTTTCAAAAAATTAATTAATTTAATTAATTTTAAATTTAAATTT 2358
QY 2644 AGCATTAAACATCATTAATAGGGAATTAATATCTATCTACTGTAATCTTTGC 2703
DB 2357 AATTTTAAATTTAAATTAATTTT-----TATTTAAATCAAAAAATTAATTAATTTTAA 2303
QY 2704 TGTATCATCTCTTAGAGATTCATATTAATTAATTAATTCATTAACATTAACGTAATG 2763
DB 2302 AAAATTAATTAATTAATTAATTAATTTTATTTTAAATTTTAAATTTAAATTA 2243
QY 2764 TCTTAGTACTGCGCATTCATAGACTGCTATTAAGCTTCACTCACTAATAC 2823
DB 2242 TTTTATTAATTAATTAATTTTATTTTATTTTATTTTCAATTTTAAATTTT 2183
QY 2824 ATGATTCACAAACCTATTTTATTCAGATGCTATTCATAAATACCTACAAATTCGG 2883
DB 2182 AATTAATAATTAATTTTATTTTATTTTATTTTAAATTTTAAATTTTAA 2123
QY 2884 TATATCTGACAAATTCCTCAAGAGCTACTTAATGGAAAGACATTTAATCTATA 2943
DB 2122 AATTAATTAATTAATTTATTAATTTATTAATTTTAAATTTTAAATTTTAAAT 2063
QY 2944 ATATAATCTCACAAACCTACAGCTTACTCTCACTTCAATTCATTTGGAAGA 3003
DB 2062 TTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT----- 2013
QY 3004 CACAAATTTCTGACATCTATATACAGAACTTACTATTTTCCCTATATATTTTCC 3063
DB 2012 AAAATTTTATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 1954
QY 3064 CTCCTGATATATCTGAGAAATCTTAATTTGGGTATCGTATGTAATCTGTTACTGC 3123
DB 1953 TTTTATTTATTTATTTAATTTTAAATTTTATTTTAAATTTTAAATTTTAAATTT 1885
QY 3124 TATTAATAACGATCATATTTGCTGCTATTAATTCAAATAGAGAGTACTATG 3183
DB 1894 AATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1835
QY 3184 AGACATATATATTTACTTCTCTGCTACCAAAATTAATTAACAGTAATTTAGTGG 3243
DB 1834 AAAATTAATTTTATTTATTTATTAATTTAAATTTAATTAATTAATTTTAAATTTT 1775
QY 3244 TAGCATATGATGATATATTAATTAATAGATATCATACCTGATTTTATATGAT 3303
DB 1774 T-----TAAATTAATAATAATAATAATAATAATAATAATAATAATAATA 1719
QY 3304 TAGCTACAGTACTTCTCTTCCAGACCTGGGAGACCAATATGATGATACAGCTCGA 3363
DB 1718 TATTTATTTATTTAAATTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTAAT 1659
QY 3364 TGTATGATCTGCTTATTAATCATCTTCAAAATTAAGTATTAACAGGGAGTTACCGAC 3423

DB 1658 TAAATTAATTTTATTTATTTTAAATTTTCAATTTATTTTCAATTTTAAATTT 1599
QY 3424 TATTTATTAATCTCAATATGCTGATTTCACTCATTTTCTCATGTAATCTACGA 3483
DB 1598 ATTAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1539
QY 3484 CTCATAGCTCTTTTACAACTTACCTTAATCTTCTGTAAGAGCTTTCTATATA 3543
DB 1538 AAAATTTTATTAATTTATTTATTTATTTATTTATTTAATTTAATTTAATTTA 1479
QY 3544 ATCAATAGCAATCCATATGAGAGACTTACAGAGTGGCAATATTAATCCATGAT 3603
DB 1478 TTTTATTAATTAATTTATTTTAAATTTTAAATTTTATTTTAAATTTTAAATTT 1419
QY 3604 CATTAATAGGCAATTAAGATACCTTCTTAAAG-----CTTTCCAGTACCT 3654
DB 1418 AAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1359
QY 3655 TAAATCATATATGATCTCTTAAACCAATATTCATTAATTAACATCAGCTCTT 3714
DB 1358 TTTAAATTTAAATTTATTTTATTTAAATTTAATTAATTTTAAATTTTAAATTT 1299
QY 3715 TTTATAGCATCTTGATGCGGATTAACCAATCTGCTAATTTAGTACATATATAC 3774
DB 1298 TTTAAATAATTTATACATTTTATTAATTAATTAATTAATTAATTAATTTTAAATTT 1239
QY 3775 ATCACTTAATATATCTTTTATTTAGATTT-CTGATATCACTTCAATCATCAACATATA 3833
DB 1238 AAAATTTATTTATTTATTTAAATTTAAATTTAAATTTAAATTTTAAATTTAAATTT 1179
QY 3834 CTTCAGAAATTAATCAATCTTCTTGACATTTCAAAATCTGTCGCAATTTACTCTA 3893
DB 1178 AATTAATTAATCAAAATTAATTAATTTTAAATTTTCAAAATCAAAATTTTAAATTT 1119
QY 3894 ATCAATATGAATTTACTGCAACAACTGACTCTTTAAAGTACCTCATATATCTACC 3953
DB 1118 CAATATTAATCACTTTTAAATCAATTTAAATTTAAATTTAAATTTTAAATTTTAA 1059
QY 3954 ACACACATTAATGATCAACACCAATTTTAAATTTCAATTTTATTTTACTTTAATTT- 4012
DB 1058 TTAATTAATTTAAATTTTCAAAATTTAAATTTAAATTTTAAATTTTAAATTTT 999
QY 4013 -----AAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 4067
DB 998 AAAATTAATTTATTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 939
QY 4068 CTATATATCAAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATA 4127
DB 938 AAAATTTTATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 879
QY 4128 ACCTATA-CTAATATTTTATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 4186
DB 878 TTTTATTAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 819
QY 4187 CATGCTATATGTTTCCATCTTAAATTTCACTGAACTTGAAGCTTTATGATTTCAATATTC 4246
DB 818 CATTTATTAATTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 759
QY 4247 AGCCAGAGCATATAGATCTATTTAGGTAATTTTAAATTTTAAATTTTAAATTTT 4306
DB 758 TAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 699
QY 4307 AATATCCACCAATGATCATCAATATGATGATGATGATGATGATGATGATGATGAT 4366
DB 698 AATTTTATTTATTTATTTATTTTATTTTATTTTAAATTTTAAATTTTAAATTTT 639
QY 4367 ATTAATCATGATCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4426
DB 638 TTTTAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 582
QY 4427 TGAAGTCAATCATCCGTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 4486

| | | | |
|----|------|---------------------------------------------------------------|------|
| D | 3267 | AATACCTAATTAAATAATGATATTTTATATATATATTAATAATTAATTATATTAT | 3326 |
| OY | 3402 | GATTAACAGGGAGTACCAGCTATTGCATATCCTCAAAAGCTGATTTCACACTCAT | 3461 |
| D | 3327 | ATATCTAATTAATTAAATBAAAAATTTTAAAATTTAAAAAGTAAATVAA-----TT | 3379 |
| OY | 3462 | TTTTCTACGTAACTCACAGACTGTATAGCTCTTTACAACCTAACCTTATCTCCCT | 3521 |
| D | 3380 | TATAAAAATTTATCTCAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTA | 3439 |
| OY | 3522 | CTGTAAAAGCTTTCTATATATCATATGAACCCATATGCAGAGACTTTACAGACGCG | 3581 |
| D | 3440 | TGATTTAATTAAATTTATATATTTATATAATTTATATATATTTATGAAATTTATATAT | 3499 |
| OY | 3582 | CATATTTACATGCCACTAATCATATATACAGGCAATAAGATATCCCTTTCTTTAAGCT | 3641 |
| D | 3500 | ATATATATATATAGAAAAATTTAAATTTATTTAAATATTTAAATTTTAAAAAT | 3559 |
| OY | 3642 | TTTCCAGTACCTTAATGCCATCAATA----TCAGATCCTTTAACATATATCCATATAT | 3697 |
| D | 3560 | TCTTAATGATATTTTATATBAAAAATTTATATBAAAAATCAAGTTTAAAAA | 3619 |
| OY | 3658 | ACAACATCAGGCTCTTTTCAATPAAGCATCTTGATCGGGATPAACCATCAACTGTAAT | 3757 |
| D | 3620 | TAAACAAAAATTTTAAATATAATTTTAAATGAATPTAATTTATTTATTTTCA | 3679 |
| OY | 3758 | TTAGGACATATATTTCACTCACTTAATATCTTTTATTAATTTGATATCAACTTCA | 3817 |
| D | 3680 | TTTTTTTAAAAATTTTAAAAAABATATTTTTTTTTTAAAAAABAACTATATCTAA | 3739 |
| OY | 3818 | TCATCAACACTAATATCTTCAAAATATATACATCTTCTTGACATTTCAAAATCCGT | 3877 |
| D | 3740 | TTATTAATTAATAGATTTATTTATATATATATAATTTAAATATTTATATATCTAAT | 3799 |
| OY | 3878 | GCAATATTAATCCTCAATCATATAGAATATCTGCAACACAA--TGATCTCTTTAAGT | 3935 |
| D | 3800 | AATTTAAATBAAAAATTTTAAAAATTTAAAAATGATATATTTATBAAAAATTTATTT | 3859 |
| OY | 3936 | ACACCTCAATATCTACACACACACACTATGATCAACACCATATATTTAATTCATTAAT | 3995 |
| D | 3860 | CTCATATTTATTTATTTATTTAATTTAATTTATATTAATATATTAATGATTTATTAATTA | 3919 |
| OY | 3996 | TTTTTACTTAATATTTTAAACATATATTAATAAATAATAGTACTATTTATTAACATTTTCA | 4055 |
| D | 3920 | TATATTTTAAAAATTTTATATATATATGAAATTTTATATATATATATATATATATGAA | 3979 |
| OY | 4056 | AGCATAGCTTCCATATATATCCAAAAAABAAAAATMAAAACAATTATBAAAAACTTATCA | 4115 |
| D | 3980 | AAATTTAATTTTAAAT-----AATTTATATTAATTTTAAAAAATTTCTTA | 4028 |
| OY | 4116 | TCAACTTTGTAAACCTATAGTACATATATTAATAATTTATTTCAATATATTAATCTTC | 4175 |
| D | 4029 | AATGATTTATTTTATBAAAAATTTTATATATATATTAATTAATTAATTAATTAATTAAT | 4088 |
| OY | 4176 | CTTGAAACAGCATGTCTATAGTTTTCATATCTAAAACTCACTTGAAGCTTTATAGATA | 4235 |
| D | 4089 | AAAAAATTTTAAATTAATTTTATATATGAAATATATATTTATTTATTTCTATTTTT | 4148 |
| OY | 4236 | TTCAATATATCAGCGCAGCAGCATATAGAAATCATATTTAGSTATTTTATATTTAAACA | 4295 |
| D | 4149 | TBAAAAAATTTTAAAAAABAAATTTTTTTTTTAAAAAABAACTATATBCTATTTATA | 4208 |
| OY | 4286 | CTAAGCAAAATATATCCCACCAATGCTATCAACATATGTGATGATGAACCATATCA | 4355 |
| D | 4309 | AATTAATATGATTTATATATATATATTAATTTATATATATTTATATATCTAATAATTT | 4268 |
| OY | 4356 | CCAACAAGAGATTTATCTACTACTCCAGTAAATAAATCTATAGTAGTTAACGGCA | 4415 |
| D | 4269 | AAATBAAAAATTTTBAAATTTAAAAATATGATATTAATTTTATBAAAAATTTATATCTCAT | 4328 |
| OY | 4416 | TTCTCAATGATGAAGTCAAAATCATCCCGATATATTCAATTTATCAACATTA-----AT | 4468 |

[illegible]

PUBMED 1455227
 REFERENCE 8 (bases 14917 to 19517)
 AUTHORS Lewis, D.L., Farr, C.L., Farquhar, A.L. and Kaguni, L.S.
 TITLE Sequence, organization, and evolution of the A+T region of
 Drosophila melanogaster mitochondrial DNA
 JOURNAL Mol. Biol. Evol. 11 (3), 523-538 (1994)
 MEDLINE 94285822
 PUBMED 8015445
 REFERENCE 9 (bases 1 to 408; 13319 to 19517)
 AUTHORS Lewis, D.L., Farr, C.L. and Kaguni, L.S.
 TITLE Drosophila melanogaster mitochondrial DNA: completion of the
 nucleotide sequence and evolutionary comparisons
 JOURNAL Insect Mol. Biol. 4 (4), 263-278 (1995)
 MEDLINE 96423163
 PUBMED 8825764
 REFERENCE 10 (bases 1 to 19517)
 AUTHORS Lewis, D.L., Farr, C.L. and Kaguni, L.S.
 TITLE Direct Submission
 JOURNAL Submitted (03-OCT-1995) Laurie S. Kaguni, Biochemistry Department,
 Michigan State University, East Lansing, MI 48824-1319, USA
 FEATURES
 SOURCE Location/Qualifiers
 1..19517
 /organism="Drosophila melanogaster"
 /organella="mitochondrion"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 /note="derived from new and previously submitted
 sequences; sequence is a composite containing sequences
 obtained from different Drosophila melanogaster strains"
 1..65
 /product="tRNA-Ile"
 complement(97..165)
 /product="tRNA-Gln"
 171..239
 /product="tRNA-Phe"
 240..1265
 /codon_start=1
 /transl_table=5
 /product="NADH dehydrogenase subunit 2"
 /protein_id="AAC47811.1"
 /db_xref="GI:1166530"
 /translation="MFNNSKILFTTMIIGLTITVTSNWLGMAGMLEINLISFIP
 LSNQNLMTSEASLEKYPLOTYASVLFSSILKMLKNNMNEINSEPTSWIMSALIL
 LKSGAPHPMPMPNMBGLTMNMLMTWOKIPMLISYINIKYLLITSVITIT
 GATGINTQTSIKIMASPSINHLGMSLSMISBSIMILFFPSPLSFVTTPMNTI
 KLFHNQPSFVPSKILKFLPMFSLGGLPFLGLPKMLVQGLTLNQFMEL
 IIMVSTLITLFFYLKICYSAPMNVENNMIMKNNMSINNMIMTFFSIFGLFL
 SLFFPMF"
 1264..1329
 /product="tRNA-Tyr"
 complement(1322..1383)
 /product="tRNA-Cys"
 complement(1403..1468)
 /product="tRNA-Tyr"
 join(1470..1472,1474..3009)
 /note="mechanism underlying reading frame shift after
 first codon uncertain"
 /codon_start=1
 /transl_table=5
 /product="cytochrome c oxidase subunit I"
 /protein_id="AAC47812.2"
 /db_xref="GI:7412849"
 /translation="MSRQWLSTYNKHDIGITYFTFGAANAVGTSLSILITAEIGHPC
 ALIGDQIVNIVTAHAFIMLFPMVPMIGFGNVLVPLMGADMAFPMNNNSFV
 LIPALSLVLSWENAGAGTGVTPPLSAGIAGASVDAIFSLHAISSILIG
 VNFITIVINMRSTGISLDRMPLFWSSVITALLILSLSPVAGAITMLTLRNTS
 FDPAGGSDPIYOHLPFPGHPEVYLLIPFGMSHIIISQSGKEKFTSGIMYAS
 LAIGLGIIVAHMFIVGMDVDPAYTSAETIAVPTGKIFSMALTGTSOYSS
 PALIAGAGVPLFTYGLGLGVLIANSSVDIILHDTYVVAHPHYLSGAVFAIMAG
 IHMYLPTGLTLNNMKSHRTIMRGVNIFFPQHFGLAGMPRYSDDYDATYNN
 IVSTIGSTLSLGLIFFFTIWMESVSRQVYFIQLNSSIEWYONTPEHRSISELF
 LLIN"
 3012..3077

```

CDS      /product="tRNA-Leu"
3083..3767
/note="tRNA scop codon is completed by the addition of 3' A
residues to the mRNA"
/codon_start=1
/transl_except=(pos:3767,aa:TERM)
/transl_table=5
/product="cytochrome c oxidase subunit II"
/protein_id="AAC47813.1"
/db_xref="GI:1166532"
/translation="MSTWALGHQDSAPLMEOLEIFEDHALLILVITVLYGLMFM
LFNNYVNPRLIGQLLEMTITLPAIILFILPSLRILYLDEINERSVTLKSGH
QWMSIYSDPFPNNIEFDSYVITPNEMLTDFRLDDNRRVLMNSGIRILVYADVI
HSMTPVPLGVKVDGTFGRLLNQTNFIFNRFELFYGCSEICGANSFMPVIESVPNV
FIMWISSNNS"
3768..3838
/product="tRNA-Lys"
3840..3906
/product="tRNA-Asp"
3907..4068
/codon_start=1
/transl_table=5
/product="ATPase 8"
/protein_id="AAC47814.1"
/db_xref="GI:1166533"
/translation="MPQMADISWLLFIPIISITFIFCSINYSYMPNSPKSNELKVI
NNSNMWKM"
4062..4736
/codon_start=1
/transl_table=5
/product="ATPase 6"
/protein_id="AAC47815.1"
/db_xref="GI:1166534"
/translation="MMNTNLFPSVDPDLAIPNFSINWLSTFGLMIPSIYWLMSRYNI
MMSILILHKRKLKLGSGHNGSTPIFSLISLNNMVGPIYFISTSLTLTL
LSALPWLICEMLYGMNTHOHPFAHLYPGCTPAIIMPEVCIETISNIRPGLAVAR
LTANMILGHLLITLIGTSSNSMYMTFLMAQIALVLESANMIQSYFAVLSLIL
YSSIVN"
4736..5524
/codon_start=1
/transl_table=5
/product="cytochrome c oxidase subunit III"
/protein_id="AAC47816.1"
/db_xref="GI:1166535"
/translation="MSTHNSPHINDYSPWPLTGATMTTYSGMWKPHQDYSLE
VLAGITLITVYQWMDVSRBGTQCHHTAVTIGLNGMILPILSVLPFSFPAAR
FHSLSFAIELGASWPWGIISFNPQIPLNTAILASGYTVTMAHSLMENNHSOT
TQGLFFTLIGIYFTILQAYEYIEAPPTIADSIVGTFPMATGFHGHVILGITFLIV
CLIRHLNHPSSKHHFGEFAAAMYHFVDVWDLFYITYYWGG"
5543..5607
/product="tRNA-Gly"
5608..5961
/codon_start=1
/transl_table=5
/product="NADH dehydrogenase subunit 3"
/protein_id="AAC47817.1"
/db_xref="GI:1166536"
/translation="MPSIIFIALILILITTVFELASILSKALIDREKSPFEGCE
PKSSRLPSLRFLFLIIFLFDVEFALILPMIIMKYNIMIMWTITSIIFILILIL
GLYENMGQAGMNSN"
5981..6045
/product="tRNA-Ala"
6055..6118
/product="tRNA-Arg"
6119..6183
/product="tRNA-Asn"
6184..6251

Query Match      3.9%; Score 184; DB 3; Length 19517;
Best Local Similarity 43.0%; Pred. No. 1,2e-14;
Matches 1844; Conservative 0; Mismatches 2355; Indels 81; Gaps 16;

```

| | | | |
|----|-------|---------------------------------------------------------------|-------|
| Db | 15162 | TATATATATATAAATAATTTTAAATTTAAATTTAAATTTTAAATTTTAACTTTTAA | 15222 |
| Qy | 353 | AGAGTCTCTTATNTGACATACATTTTAAATTCATCCATCAACAAATAATGTACTATCTA | 412 |
| Db | 15222 | AAATATTTTTTTTTTAAABAAAAATTTTTTTTTTAAABAAAAATTTTTTTTTTAAABAAAT | 15281 |
| Qy | 413 | ATTATATAAAAAATCCCATAGTCTTTAGCAGAACTACAAAATATATAGACACTTAA | 472 |
| Db | 15282 | TTTAAAAATATAGATTAATTTCTTTTAAATGACTAAAAAAATTTTTTTTTTAAGTAT | 15341 |
| Qy | 473 | CTTATTA--ATAATTTCAAAATTTATATATCTACTCATATTTTGCAATAATTCACAT | 530 |
| Db | 15342 | TTTAAACCTTTTTTTTTCACATTTTTTAAABAAATATATATATATTTTAAABAAATTT | 15401 |
| Qy | 531 | CTGACATGATTTAACATATGTATACACATTTCTTAACTATGATATATACATTAATATAT | 590 |
| Db | 15402 | TTTTTTTTTAAAAAAAATGAAAAATTAATATATAAAAATTTTTTTTTTACAAAAATGAAA | 15461 |
| Qy | 591 | TATATTTTTCACGGTTCTTGATATATATATATAGTAAATAAGAAATCAGTATTTTG | 650 |
| Db | 15462 | TTTATTCATTTAAAAAAAATTTATTTAAATTTTTTATGATATATATATATATATATAT | 15521 |
| Qy | 651 | GTGACAGTCAATTTGGACACAGCAATAGCAATTGCACTGCAGCACATGTATATCAGTAA | 710 |
| Db | 15522 | TTATTTAAAAATCAATATATATATATATATATATATATATATATATATATATATATAT | 15581 |
| Qy | 711 | ACTATGCGGCGATGATCATAGAAATTTACACATATAAACCTTACCGAAAAATTTTA | 770 |
| Db | 15582 | TATATATATATATATATATATATATATATATATATATATATATATATATATATATAT | 15641 |
| Qy | 771 | AAATATTTACCAATATCATCTACACAGCAACATATATGCAACAGCAATATAGACGAG | 830 |
| Db | 15642 | TATATATATATATATATATATATATATATATATATATATATATATATATATATATAT | 15696 |
| Qy | 831 | TATATATGCAACAAATACATATATATCTTACATATCTTACACACATATACGACA | 890 |
| Db | 15697 | AAATTTATTCCTTATTCATTAATTTATTTATATATATATATATATATATATATATAT | 15756 |
| Qy | 891 | TATGTACAAATACAAACAGCAACAGCATATGTATAAAAATCTCCATATTTAATTTGTA | 950 |
| Db | 15757 | TTAAAAAAAATTTTATTTAAATTAATCTAATATTAACATTTTATTAATTAATTTATTT | 15816 |
| Qy | 951 | GTAAGGATGGAATATACATCTCCAAATTTCCCGTGAATAGAGAAAGAAATTTTAC | 1010 |
| Db | 15817 | ATATATATATATATATATATATATATATATATATATATATATATATATATATATAT | 15876 |
| Qy | 1011 | AAATATATCCAAATTTTATATCTCTGTGTCGAAGTTTGCTAAGAAATGGACAAATC | 1070 |
| Db | 15877 | ATATATTTTAAAAAATTCATTTATATATTTTATATATATATATATATATATATATAT | 15936 |
| Qy | 1071 | TTCTCTAGATATGTCTGTGCGGATATATTAAGAATCTGTGTGAATCATGTATAGAA | 1130 |
| Db | 15937 | TCAATTTAAATTTATATATATATATATATATATATATATATATATATATATATATAT | 15996 |
| Qy | 1131 | CAATAGTAATGATGTTCTTAAAAATATATATCCATCAAGATAATTAGGTGTACAGATTG | 1190 |
| Db | 15997 | TAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT | 16051 |
| Qy | 1191 | GAGGTGATTTAAAGACATATTTGCAATGCAATGCAATGCAATATCGGTGAAAAAATTTAG | 1250 |
| Db | 16052 | TTTCAATATATTTATGTATATATTTAAACCTTAAAAATATTTTTTTTTTAAABAAAAATGA | 16110 |
| Qy | 1251 | GTAATTAATGCTGTGCTACTGTTATATCTTAAGGCAATGATGAATAATTAACACTATATAT | 1310 |
| Db | 16111 | TTTATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT | 16170 |
| Qy | 1311 | TAGCAAAAAATTCATTCATATGATCTTCATTCATTAATTTGTGCATCATGCTTGAGATC | 1370 |
| Db | 16171 | TTTAAATATATATATATATATATATATATATATATATATATATATATATATATATAT | 16230 |
| Qy | 1371 | TATATTTAACATGTACACAGAAATTCACGCAATATGCTTTTGGACTAGAAATAGGAA | 1430 |

[illegible]

QY 2502 ATTACGCTGTAATGCTGAACCTTCATTCCAAACAAATTCAGACAGTGTGTATACATTAC 2561
 Db 17349 AACTATATATATATATATATATATATATATATATATATATATATATATATATATATAT 17408
 QY 2562 CTCCTCAAAAGAGTAACGAGTGTATATATATATATATATATATATATATATATATAT 2621
 Db 17409 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 17468
 QY 2622 GATATGCTACAGAAATPACTTGAACCTTAAACATCATTAATPAGCGAATPAGTACTA 2681
 Db 17469 TATTGAAATTTTATATATATATATATATATATATATATATATATATATATATAT 17528
 QY 2682 TATCTACGTGTAATCTTTTGTGCTATATCTCTTGTAGAGATTTCTATATATATATCC 2741
 Db 17529 TATATATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 17588
 QY 2742 ATTCTATTAACATTTACGTAATGTCTTAAAGTACCTGCGCATTCATATGACATTCGTA 2801
 Db 17589 AATCATGTTTTTTTTAAAAAATTAACAAAAATTTTAAATTAATATTTTAAATGAAATA 17648
 QY 2802 TTTAAGCTTCATCACTTAATPACATGATGACAAACCTATTTTTTAAAGATGCTATTC 2861
 Db 17649 TATTTATTTATTT-----TTTCATTTTTTAAAAAATTTTTTAAAAAATTAATTTTT 17703
 QY 2862 TAAATATACCTACAAATTCGCTATATCTGTACATATTTCTACCTAAGACGCTACTCA 2921
 Db 17704 TTTTAAAAAATTAATATATATATATATATATATATATATATATATATATATATAT 17763
 QY 2922 TTGGAAGGACATTTATCTATATATATATATATATATATATATATATATATATATAT 2981
 Db 17764 TATATATATATATATATATATATATATATATATATATATATATATATATATATAT 17823
 QY 2982 TTTCATATCTTTGGAAGAGACACATATATCTCACATTAATCTACAGAAATCTTAC 3041
 Db 17824 AATATATTTTAAAAAATTTATTTCTCATATTTATTTATTTATTTATTTATTTATTTAA 17883
 QY 3042 TATTTTCCCTATATATTTTCCCTCGTATATCTGATATCTGATATCTTATTTGCTATCTG 3101
 Db 17884 TATATATATATATATATATATATATATATATATATATATATATATATATATATAT 17942
 QY 3102 ATCGTAATCTGTATCTCATATATATATATATATATATATATATATATATATATATAT 3161
 Db 17943 AATATATATATATATATATATATATATATATATATATATATATATATATATATAT 18002
 QY 3162 CAATATGAGAGGATCTGTAGAGACAAATATATATATATATATATATATATATATAT 3221
 Db 18003 TTTAAAAATTTCTTAATGATATATTTTAAAAAATATTTATATATATATATATATAT 18062
 QY 3222 TAAACAAGTAATTTAGCTGTAGATAGATAGATAGATAGATAGATAGATAGATAGAT 3281
 Db 18063 TTTAAAAAATTAACAAAAATTTTAAATTAATTAATTAATTAATTAATTAATTTT 18122
 QY 3282 ACCCTGTGATTTTTATGATATTTAGCTGACGAACTACTCTTTTCAACACCTGGCGAAC 3341
 Db 18123 ATTTTCAATTTTTTTAAAAAATTTTTTAAAAAATATTTTTTTTTTAAAAAATCT 18182
 QY 3342 CAGTATAGATATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGAT 3401
 Db 18183 AATATCTATATATATATATATATATATATATATATATATATATATATATATATATAT 18242
 QY 3402 GTATTAACAGGGAGTACCGACTATTTCTATATCTCTCAATAGCTGATTCACATCAT 3461
 Db 18243 AATATCTATATATATATATATATATATATATATATATATATATATATATATATAT 18295
 QY 3462 TTTCTCTAGATATACCAAGCTATAGCTCTCTTCACTAATCTTATATCTTCTT 3521
 Db 18296 TATATATATATATATATATATATATATATATATATATATATATATATATATATAT 18355
 QY 3522 CTGTAAAAAGCTTTTCTATATATATATATATATATATATATATATATATATATATAT 3581
 Db 18356 TGAATTAATTAATTTATATATATATATATATATATATATATATATATATATATATAT 18415

QY 3582 CAATATPACCATGCCCACTAATCATATATACAGGCAATATAGATATCTTTCTTTAAGCT 3641
 Db 18416 AATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 18475
 QY 3642 TTTCCAGTACCTTAATPCCATCAATA-----TGAGATCCCTTAAACATATATCCATAT 3697
 Db 18476 TCTTAAATGATATATTTTTTTAAAAAATTTTATATATATATATATATATATATATATAT 18535
 QY 3698 ACAACATCAGGCTCTTTTCTATAGGACATTTGATGCGGATTAACCATCAATGCTAAT 3757
 Db 18536 TAAACAAAAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTTTCA 18595
 QY 3758 TTAGGACATATATATATATATATATATATATATATATATATATATATATATATAT 3817
 Db 18596 TTTTTTAAAAAATTTTTTAAAAAATTAATTTTTTTTTTTTTTAAAAAATCTATATCTTA 18655
 QY 3818 TCATCAACACTAATCTTCAAGAAATATACAAATCTTCCCTGGACATTTCAAAATCTGT 3877
 Db 18656 TATATATATATATATATATATATATATATATATATATATATATATATATATATAT 18715
 QY 3878 GCCATATTTTACCTACATATCTAATGAAATATCTGCAACCA--TGTACTCCTTTAAGT 3935
 Db 18716 AATTTAAATTAATTAATTTTAAAAATTTTAAAAATGATATATATATATATATATAT 18775
 QY 3936 ACACCTCATATATCTACACACACACTAATGATCAACCAATATATATATATATATAT 3995
 Db 18776 CTATATATATATATATATATATATATATATATATATATATATATATATATATAT 18835
 QY 3996 TTTTCTATATATATATATATATATATATATATATATATATATATATATATATATAT 4055
 Db 18836 TATATATATATATATATATATATATATATATATATATATATATATATATATATAT 18895
 QY 4056 AGCATAGGTTTCTATATATATATATATATATATATATATATATATATATATATATAT 4115
 Db 18896 AATTTAATTTATTTAAT-----AATTTAATTAATTTTAAAAATTTCTTA 18944
 QY 4116 TCACCTTGTAAACCTATATAGTATATATATATATATATATATATATATATATATAT 4175
 Db 18945 AATGATATTTTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTAAC 19004
 QY 4176 CCGTGAACAGGATCTATATATATATATATATATATATATATATATATATATATATAT 4235
 Db 19005 AAAAAATTTTATATATATATATATATATATATATATATATATATATATATATAT 19064
 QY 4236 TTCAATATTCAGCCGACAGCATATAGATATATATATATATATATATATATATATAT 4295
 Db 19065 TAAAAAATTTTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAA 19124
 QY 4296 CTAGCAAAATATATATATATATATATATATATATATATATATATATATATATATAT 4355
 Db 19125 AATTAATATATATATATATATATATATATATATATATATATATATATATATATAT 19184
 QY 4356 CCACAAGAGTATATATATATATATATATATATATATATATATATATATATATATAT 4415
 Db 19185 AATATAAAAATTTTAAAAAATTTTAAAAAATATAGATATATATATATATATATAT 19244
 QY 4416 TCTCCATATATATATATATATATATATATATATATATATATATATATATATATAT 4468
 Db 19245 ATTTATTTATTTATATATATATATATATATATATATATATATATATATATATAT 19304
 QY 4469 ATGTCCAAACATATATATATATATATATATATATATATATATATATATATATATAT 4528
 Db 19305 ATTTATATATATATATATATATATATATATATATATATATATATATATATATAT 19364
 QY 4529 ATATTTCTGAACTGTTATCTGTAAGTAT 4588
 Db 19365 AATATATTTAAATTAATTTTTCATTAATTT 19394

RESULT 18
 AC093899/c 172816 bp DNA linear PRI 12-JUN-2002
 LOCUS
 DEFINITION Homo sapiens BAC clone Rpi1-724016 from 2, complete sequence.

ACCESSION AC093899 AC068884
 VERSION AC093899.3 GI:18497265
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS Sulston, J.E. and Waterston, R.
 TITLE Toward a complete human genome sequence
 JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
 MEDLINE 99063792
 PUBMED 9847074

REFERENCE
 AUTHORS Pearman, C., Haakenson, W. and Boyer, E.
 TITLE The sequence of Homo sapiens BAC clone RP11-724016
 JOURNAL Unpublished (2001)
 REFERENCE 3 (bases 1 to 172816)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (10-SEP-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 4 (bases 1 to 172816)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (05-FEB-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 5 (bases 1 to 172816)
 AUTHORS Waterston, R.
 TITLE Direct Submission
 JOURNAL Submitted (12-JUN-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Feb 5, 2002 this sequence version replaced gi:115625013.

REFERENCE
 AUTHORS Center: Washington University Genome Sequencing Center
 CENTER Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc
 Contact: sapiens@wustl.wustl.edu
 Summary Statistics
 Center project name: H_NH0724016
 Drafting Center: WIBR

COMMENT
 NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
 Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
 The RP11-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Frengen, E., Taten, M., Catanesi, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong and coworkers at http://www.chori.org
 VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:
 Actual start of this clone is at base position 1 of RP11-724016;
 actual end is at base position 172816 of RP11-724016.
 Unresolved simple sequence repeat from base position 22700 to 25900.

FEATURES
 source The sequence of AC068884 has been incorporated into AC093899.
 location/Qualifiers

1..172816
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="2"
 /map="2"
 /clone_id="RP11-724016"
 /clone_idb="RP11-724016"
 19..1140
 /rpt_family="L1"
 1141..1437
 /rpt_family="Alu"
 1438..2657
 /rpt_family="L1"
 2658..2729
 /rpt_family="L1"
 2730..3320
 /rpt_family="L1"
 3321..3642
 /rpt_family="Alu"
 3643..4184
 /rpt_family="L1"
 4185..4996
 /rpt_family="ERV1"
 5011..5263
 /rpt_family="L1"
 5285..5907
 /rpt_family="CR1"
 6895..7078
 /rpt_family="L1"
 7103..7421
 /rpt_family="Alu"
 7807..7952
 /rpt_family="L1"
 7981..8289
 /rpt_family="Alu"
 8360..8551
 /rpt_family="L1"
 8622..8718
 /rpt_family="MIR"
 8784..8825
 /rpt_family="(TTTC)n"
 8840..9134
 /rpt_family="Alu"
 9136..9174
 /rpt_family="Alu"
 9861..9881
 /rpt_family="(A)n"
 9882..10166
 /rpt_family="Alu"
 16502..16556
 /rpt_family="L2"
 16866..17165
 /rpt_family="L1"
 17179..17482
 /rpt_family="Alu"
 18025..18119
 /rpt_family="MIR"
 18177..18471
 /rpt_family="Alu"
 19312..19441
 /rpt_family="MER1_type"
 19731..20040

```
repeat_region /rpt_family="MaLR"
20403. 20472
/rpt_family="L2"
repeat_region /rpt_family="(CA)n"
21286. 21344
22724. 23977
/rpt_family="(TA)n"
repeat_region 24005. 24171
/rpt_family="(TA)n"
repeat_region 24196. 25007
/rpt_family="(TA)n"
repeat_region 25030. 25205
/rpt_family="(TA)n"
repeat_region 25223. 25398
/rpt_family="(TA)n"
repeat_region 25395. 25445
/rpt_family="AT_rich"
25456. 25919
/rpt_family="(TA)n"
repeat_region 25920. 26206
/rpt_family="Alu"
repeat_region 26893. 26917
/rpt_family="AT_rich"
27194. 27497
/rpt_family="L1"
repeat_region 27498. 27814
/rpt_family="MER2_type"
repeat_region 27815. 28151
/rpt_family="L1"
repeat_region 28152. 28278
/rpt_family="Alu"
repeat_region 28279. 28485
/rpt_family="L1"
repeat_region 28795. 28829
/rpt_family="AT_rich"
repeat_region 28986. 29100
/rpt_family="L2"
repeat_region 29237. 29349
/rpt_family="L1"
repeat_region 29350. 29484
/rpt_family="Alu"
repeat_region 29485. 29589
/rpt_family="L1"
repeat_region 29811. 29874
/rpt_family="L2"
repeat_region 29875. 30284
/rpt_family="L1"
repeat_region 30320. 30522
/rpt_family="L1"
repeat_region 30523. 30647
/rpt_family="Alu"
repeat_region 30848. 30882
```

[illegible]

| | | | |
|----|-------|--------------------------------------------------------------------|-------|
| QY | 538 | GATTACCAATGTGTACACCATCTTCTTAACAGTAGTAATATACAAATTAATATATAT | 597 |
| Db | 25394 | TATATAAATATATATAATGTATATATATAATATATATATATATATAATATAAT | 2553 |
| QY | 508 | TTTGACGTGGTTCTTGA---TATATAATGTGATAAAAAATGAAATACGATTTTGGTGC | 654 |
| Db | 25534 | ATATATATATATATATATATGTATATATAATATATATAATTTATATAATAAAGTATATATA | 25477 |
| QY | 655 | AGATATCATTTGGCAGACAAATAGCAATTGCACTGTCCAGACATGTGTATACATTAACCT | 714 |
| Db | 25474 | ATATATATTTTATATATATATAATATATATAATTTATATATAATATATATAAAAAATC | 25411 |
| QY | 715 | ATGGGACGTGATCATTGAAATATTACATATPAAACCTTACCGAAAAATTTAAAAAT | 774 |
| Db | 25414 | TATATATATTTATATATAATTTATATATATATATATATCTTTTATATATATTAATATATATAT | 25355 |
| QY | 775 | TTTACCCGATATCATCTACACAGCAATATATGTGCAACGAGCAATATAGACAGATAT | 834 |
| Db | 25354 | TAAATATATAATATATATTAACGATATAATATATATATATATATATATATTAACATATAAAT | 25292 |
| QY | 835 | ATCTGACACATATACATGATATATCTTAAGTATCTTACACAACTTACGACCATATG | 894 |
| Db | 25294 | ATATATATTAACATATAATATATATATATATATAATATATATAACATATAAATATTTATAT | 25233 |
| QY | 895 | TACACAAATACACACAAACAGCATATGTATAAATACTCCATATTTATTTGTATGTA | 954 |
| Db | 25234 | ATATAATATATTTAAACATATAATATATATATATATATATATATATATTAACATATAAT | 25177 |
| QY | 955 | AGGATATGMAATTATCATCTCAATTTCCCACTGAAATGCGAAGAAATTTTACATAT | 1014 |
| Db | 25174 | ATTATATATATAATATATATATATAACGTAATATATTTATATAGATATAATATATAACGT | 25111 |
| QY | 1015 | TATTCACATTTTATATCTCTCTGTCGACAGTTTGTCTAAAGAAATGACAGACATCTCC | 1074 |
| Db | 25114 | AAATATATATTTATCTATATAATATATATATTAAGCTAAATATATATTTATCTATATAATATAT | 25055 |
| QY | 1075 | TTTGATGATATGACTCTGCTGTGATATATAAGACTTGGTGATCATTTGATAGAAACAT | 1134 |
| Db | 25054 | TATTTAAACATPAAATATATATTTTATAGTAAATATATATTAAGTAAATATATAATTTATC | 24995 |
| QY | 1135 | AAGTATGATGTTCTAAAAATATATATACATCCAGATATTTATGTGTGACAGATTGGAC | 1194 |
| Db | 24994 | TATATAATATATATAACGTAAATATATATTTATCTATAACATATATCAATATATAATAT | 24933 |
| QY | 1195 | TGCAATTAAAGACATAATTGCATTGCAATGTGATGTAATAATCGCTGGAATAATTTAGCTAT | 1254 |
| Db | 24934 | TATTAATATATATAATAATATATATATCAAAAGTAAATATATTAATATATATAAATATATAT | 24875 |
| QY | 1255 | TATGCTGTGTGCTACTGTTTAACTAAAGCATGATGAATAATTAACAATATATCTATAC | 1314 |
| Db | 24874 | CAATGTAAATATATATATATATATATATAATATAATATCAATGTGAAATATATTAATATATA | 24815 |
| QY | 1315 | AAAAAATCATTCAATGATCTTATATCATCTTAATGGTCCATCATGTCTTGAGATCTAT | 1374 |
| Db | 24814 | TAAATATATATACAAATATATAATATATATATATCAAAATATATAATATATATTT---ATATGTAA | 24755 |
| QY | 1375 | ATTAAACATGTACAACAACTTACGCAATATGCGCTTTTGGCTAGTAATAGGAAAAAG | 1434 |
| Db | 24757 | TATATATTAATATATAATATATATATATGTATATATATAACATATATATTTAAATATATATA | 24698 |
| QY | 1435 | TAGAAATATAATATACATTAATAGATACAAACCTTAAGCTTTGTGAAGAACAGTACTGT | 1494 |
| Db | 24697 | TTTATATATATAACATTTAATATCA---TATATATTTATATATATAATATATTTTA | 24633 |
| QY | 1495 | AAAAACATGATATCATTTGCAAAAAAACTTAATGTAGAACTACCAATTTGCATATCTAT | 1554 |
| Db | 24642 | ATATCAATATATATTTTATATATATAATATTTAAATATAAATAATATTTTATATATAATAT | 24583 |
| QY | 1555 | TTTCAATTTTATTCATGAGAAATATATCATGTAGTAAAGCATATCAAACTATTTATCTTA | 1614 |
| Db | 24582 | ATATTAATATCAAAATGTATATAATATATATATTTTATATAATGTATATAAACATATAAT | 24523 |
| QY | 1615 | GTCTATCATCTTTGATTAATTTCAACACATGATATAATTAACGTCAAAATATAGAAAT | 1674 |

| | | | | | | | |
|----|-------|-------------|------------------------------------|------------------------------------------------|----------------------|---------|------|
| D | 24522 | TATTTATATAA | CTGTATTTAAATATATTAATATATTTATATATTA | CTGTATTTAAATATATTAATATAT | 24463 | | |
| OY | 1675 | CAAGCTT | CACACAGATATATTA | AAAGCAGTCAAAATGCCATTCGTAAATCCATCACTA | 1734 | | |
| D | 24462 | TTTATATTA | CTGTATTTAAATATTAATATATTTATATTA | CTGTATTTAAATATTAATATAT | 24403 | | |
| OY | 1735 | TGTAA | CATATTACATTA | CAAACTTAACTATACCTTAAATTTCTGACTATATATTA | 1794 | | |
| D | 24402 | TATATTA | CTGTATTTAAATATATTAATATATTTATATTA | CTGTATTTAAATATATTAATATAT | 24343 | | |
| OY | 1795 | CAAAATTC | CAATCTCTATTCATTA | CAATTAACCAATTTGTTATCGATTTATTAATTAAGTCA | 1854 | | |
| D | 24342 | TATTA | CTGTATTTAAATATATTTATATTA | CTGTATTTAAATATTAATATAT | 24286 | | |
| OY | 1855 | AAAC | TTTACTTATTTATTA | AAATATATTTGTTGACACTTTTCAGTACATCATGTATAT | 1914 | | |
| D | 24285 | ATATTA | CTGTATTTAAATATATTAATATTTATATTA | CTGTATTTAAATATTAATATAT | 24222 | | |
| OY | 1915 | AATGA | CATATTTTACATTA | AAATTAATTAATTA | 1974 | | |
| D | 24225 | AACGTAT | TAATAATATATTTATATTA | CTGTATTTAAATATTAATATAT | 24166 | | |
| OY | 1975 | TAAAC | ACCTATTTACTACGTAAATATATTA | CAATGATGCTGACATATTAATCTATCAT | 2034 | | |
| D | 24165 | AAATAT | TAATAATATTAATATTTATTA | TAATATTAATTAATATTAATATTA | 24106 | | |
| OY | 2035 | CCTTG | ACATTTTCAAACTATTA | AAATATGATTAACGTTTTTTATGCCCATTA | 2094 | | |
| D | 24105 | TATATTA | ATATATTAATATTAATATTAATATTA | TAATATATTTATATTAATATTAAT | 24046 | | |
| OY | 2095 | ATATAA | ATTCGTAATTA | CCCAAAAGTTAAATATACATTCCTGACACATATGAAGA | 2154 | | |
| D | 24045 | ATTAAT | ATATTAATATTAATATTAATATTAATATTA | TAATATTAATATTAATATTAAT | 23989 | | |
| OY | 2155 | ATACA | TATCA | CGTCAATTAATTAACAAATTTCTAATATTAACACCTACTTAACATA | 2214 | | |
| D | 23988 | ATTAA | AAATATTAATATTAATATTAATATTA | TAATATTAATATTAATATTAATATTA | 23929 | | |
| OY | 2215 | GGAAC | CAATGATATCATACCTA | AGTAAGATATCCCTGACAGATATCAAAATATA | 2274 | | |
| D | 23928 | ATATAT | TAATTAATATTAATATTAATATTAATATTA | TAATATTAATATTAATATTAAT | 23869 | | |
| OY | 2275 | GAAAA | ATCGTATATTAATTAATTA | TACAATTAACAATTTCCAACTTCACACAAATCTCA | 2334 | | |
| D | 23868 | TATATA | TAATATGTTTATATTAATTAATTA | ATTAATATTAATATTAATATTAAT | 23809 | | |
| OY | 2335 | AAGT | CTATATTAATTA | CAAAATTCAGAAATTAATATTA | CTTCAACACATATTAAT | 2394 | |
| D | 23808 | TATATAT | TAATTAATTAATTAATATTAATTAATTA | TAATTAATTAATTAATTAAT | 23749 | | |
| OY | 2395 | TCTAT | ATATACAT | CATATCCTTTTAAAAAGATCTTCA | CAATCATCAATTAAT | 2453 | |
| D | 23748 | TATAT | AAAAATTAATTTATATTAATTAATTA | ATTAATATATGATATTAATTAAT | 23689 | | |
| OY | 2454 | ATACA | TGCAATGTTTATTC | GAAATTCATACAAATTCAGAAATTTTCAATTA | CGGTGA | 2513 | |
| D | 23688 | ATATA | TGTAATATTAATTAATTAATTAATTA | TATATTAATTAATTAATTAAT | 23629 | | |
| OY | 2514 | ATGCT | GACCTTCATTTCC | CAACAAATTCAGACAGTGTGATACATTA | CTCCAAACGAG | 2573 | |
| D | 23628 | ATATTA | ATTAATTAATTAATTAATTAATTAATTA | TAATTAATTAATTAATTAAT | 23569 | | |
| OY | 2574 | ATTA | CTGAGATTTTAAAT | ACTGCTTCAAAATTTCTTCA | GAGTTTACGTAA | TGGTACG | 2633 |
| D | 23568 | TATTA | ATATTTAAATTAATTAATTAATTAATTA | TAATTAATTAATTAATTAAT | 23509 | | |
| OY | 2634 | AAAT | TAATCTTACGACTTAA | AAACATCATTAATATAGGCAATTA | AGTACTATATCTACGTGA | 2693 | |
| D | 23508 | ATTTA | TAATATTAATTAATTAATTAATTAATTA | TAATTAATTAATTAATTAAT | 23449 | | |
| OY | 2694 | AACTTT | TGCTGTAATCATCT | CCTTGAGAGATTTCA | TAATTAATTCATTTAATCAT | 2753 | |

| | | | |
|----|-------|-------------------------------------------------------------------|-------|
| Db | 23448 | ATATATGTTATATATATATAT-----ATATTTATATATATATATATATATATATATA | 23399 |
| Qy | 2754 | TACGTAATATGCTTTAGTACACCTGGCAATTCATATACATGCAATGGTATATTAAGCTTCAT | 2813 |
| Db | 23396 | TATATGTTTATATATATATATATATATATATATATATATATATATATATATATATATATGTTTAT | 23337 |
| Qy | 2814 | CACCTTAATACATGATPACAAACCTATTTTTTTTACAGATGCTATTCATTAATAATACCTAC | 2873 |
| Db | 23336 | ATATATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT | 23277 |
| Qy | 2874 | ACAAATCCGGTATATCTGTCAATATTTCTACATTAAGCGTACTCTAATTTGGAGACAT | 2933 |
| Db | 23876 | ATGTTTATATATATATATATATATATATATATATATATATATATATATATATATATATATAT | 23217 |
| Qy | 2934 | TTA-----ATCTATATATATTAATCTCTCAAAACCTACACGCTTTTACTTCACTTCAAT | 2988 |
| Db | 23216 | ATTAATATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT | 23157 |
| Qy | 2989 | ATCTTTGGAAGACACATATATTTCTCAATCTATCTACACAGATCTTACTATTTTC | 3048 |
| Db | 23156 | ATTAATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT | 23097 |
| Qy | 3049 | CTATATATTTT 3061 | |
| Db | 23096 | TAAATATTAATAT 23084 | |

| | | | | | |
|------------|-----------------------------------------------------------------------------------|------------|-----|--------|-----------------|
| RESULT 19 | | | | | |
| AE001398/c | | | | | |
| LOCUS | AE001398 | 14867 bp | DNA | linear | INV 04-OCT-2002 |
| DEFINITION | Plasmodium falciparum 3D7 chromosome 2 section 35 of 73 of the complete sequence. | | | | |
| ACCESSION | AE001398 | AE001362 | | | |
| VERSION | AE001398.1 | GI:3845197 | | | |
| KEYWORDS | | | | | |
| SOURCE | Plasmodium falciparum 3D7 | | | | |

| ORGANISM | REFERENCE |
|-------------------------------------------------------------|-----------|
| Plasmodium falciparum | 357 |
| Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium | |
| 1 (bases 1 to 14867) | |

AUTHORS
Gardner, M.J., Tettehin, H., Carnucci, D.U., Cummings, L.M., Ariving, L., Koonin, E.V., Shallow, S., Mason, T., Yu, K., Fujii, C., Pederson, J., Shen, K., Jing, J., Aston, C., Lai, Z., Schwartz, D.C., Pereira, M., Salzberg, S., Zhou, L., Sutton, G.G., Clayton, R., White, O., Smith, H.O., Fraser, C.M., Adams, M.D., Venter, J.C. and Hoffman, S.L.

TITLE
Chromosome 2 sequence of the human malaria parasite *Plasmodium falciparum* (2001) 1102-1110 (1999)

UOORNAU
SCIENCE 282 (1991), 1140-1152 (1996)
MEDLINE
99021743
PUBMED
9804551
2 (bases 1 to 14867)
REFERENCE
Gardner, M.J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R.W., et al. 1996. The genome sequence of *Escherichia coli* O157:H7. *Science* 272: 205-208.

TITLE Genome sequence of the human malaria parasite *Plasmodium falciparum*
JOURNAL Nature 419, 498-511 (2002)
REFERENCE 3 (pages 1 to 14867)
AUTHORS Gardner, M.J.,
TITLE Direct Submission
JOURNAL Submitted (02-OCT-2002) The Institute for Genomic Research, 9712
 Medical Center Dr, Rockville, MD 20850, USA
 Location/Qualifiers

```
source
1.14867
/organism="Plasmodium falciparum 3D7"
/mol_type="genomic DNA"
/isolate="3D7"
/db_xref="taxon:36329"
```

| | | | |
|---------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|-------------------------------------------------------------------------------------------------------------------------------------|
| repeat_region | /chromosome="2" complement(<1. .20) /note="AT rich" | repeat_region | complement(2446. .2468) /rpt_type=tandem /rpt_unit="(TA)n" |
| repeat_region | /rpt_type=tandem complement(65. .137) /note="AT rich" | repeat_region | /rpt_type=tandem /rpt_unit="(TA)n" |
| repeat_region | /rpt_type=tandem complement(1178. .211) /rpt_type=tandem /rpt_unit="(TA)n" | repeat_region | /rpt_type=tandem complement(2693. .2736) /rpt_type=tandem /rpt_unit="(TA)n" |
| repeat_region | complement(307. .327) /rpt_type=tandem /rpt_unit="(A)n" | repeat_region | complement(2697. .2732) /rpt_type=tandem /rpt_unit="(TA)n" |
| repeat_region | /rpt_type=tandem complement(328. .348) /rpt_type=tandem /rpt_unit="(TA)n" | repeat_region | complement(2757. .2845) /note="AT rich" |
| repeat_region | /rpt_type=tandem complement(378. .423) /note="AT rich" | repeat_region | /rpt_type=tandem complement(2856. .2880) /rpt_type=tandem /rpt_unit="(TA)n" |
| repeat_region | /rpt_type=tandem complement(479. .503) /note="AT rich" | repeat_region | complement(2884. .2953) /rpt_type=tandem /rpt_unit="(TA)n" |
| repeat_region | /rpt_type=tandem complement(513. .539) /rpt_type=tandem /rpt_unit="(A)n" | repeat_region | /rpt_type=tandem /rpt_unit="(TA)AA)n" |
| repeat_region | 577. .633 /rpt_type=tandem /rpt_unit="(A)n" | repeat_region | complement(2924. .2988) /rpt_type=tandem complement(2977. .3167) /rpt_type=tandem /rpt_unit="(TAA)n" |
| repeat_region | /rpt_type=tandem complement(714. .747) /note="AT rich" | repeat_region | complement(3047. .3132) /rpt_type=tandem complement(3047. .3132) /rpt_type=tandem /rpt_unit="(TA)n" |
| repeat_region | /rpt_type=tandem complement(888. .918) /note="AT rich" | repeat_region | complement(3352. .3514) /rpt_type=tandem /rpt_unit="(TA)n" |
| repeat_region | /rpt_type=tandem complement(1006. .1049) /note="AT rich" | repeat_region | complement(3532. .3559) /note="AT rich" |
| repeat_region | /rpt_type=tandem complement(1075. .1120) /note="AT rich" | repeat_region | /rpt_type=tandem complement(3689. .3720) /note="AT rich" |
| repeat_region | /rpt_type=tandem complement(1140. .1283) /note="AT rich" | repeat_region | /rpt_type=tandem complement(4024. .4114) /note="AT rich" |
| repeat_region | /rpt_type=tandem complement(1294. .1346) /rpt_type=tandem 1298. .1346 | repeat_region | /rpt_type=tandem complement(4107. .4146) /rpt_type=tandem complement(4115. .4146) /rpt_type=tandem /rpt_unit="(TA)n" |
| repeat_region | /rpt_type=tandem complement(1423. .1531) /note="AT rich" | repeat_region | complement(4227. .4257) /rpt_type=tandem /rpt_unit="(TA)n" |
| gene | /rpt_type=tandem complement(<1570. .>2424) /locus_tag="PEB0490c" | repeat_region | /rpt_type=tandem complement(4338. .4386) /note="AT rich" |
| mRNA | complement(<1570. .>2424) /locus_tag="PEB0490c" | repeat_region | /rpt_type=tandem complement(4403. .4450) /note="AT rich" |
| CDS | complement(1570. .2424) /locus_tag="PEB0490c" | repeat_region | /rpt_type=tandem complement(4545. .4785) /note="AT rich" |
| | /codon_start=1 /product="hypothetical protein" | repeat_region | /rpt_type=tandem complement(4937. .5001) /note="AT rich" |
| | /protein_id="AAC71887.1" | repeat_region | complement(5102. .5191) /note="AT rich" |
| | /db_xref="GI:3845198" | repeat_region | /rpt_type=tandem complement(5246. .5275) |
| | /translation="WKEKRSKIMDYVVDVREKSGKNSLLKSSSTKSDYKK SISIFPKSDSHKSGSFKGRSGIFINRSKSVKSGYNNLLINKNNVYGNFHNGR DNFPRKSGFGRVFDNRKSGFKRFLSNRKSGVKSRYRNGSNMGKRSFNAPR TVVTRILNNYKVASPVKKFNNNLSLRKRRRTALNTKSGPVGITKSSVPKRIKK GLKSKSKRKRSKSGSKFKPLNNKYFLSKIKITSLNKIPSPLEKGRNTEVNLPSLS NNATKKKK" | | |

QY 2246 TATCCCTACAGATGATCAATATATAAGAAAATCGTATGATATTTATACATTTAC 2305
 DB 7866 TTTTATATATATATATATATATATATATATATATATATATATATATATATAT 7807
 QY 2306 AAAATTTCCATCTCAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAA 2365
 DB 7806 TATTTATATATATATATATATATATATATATATATATATATATATATATATAT 7747
 QY 2366 TATTTATATATATATATATATATATATATATATATATATATATATATATATAT 2425
 DB 7746 TATTTATATATATATATATATATATATATATATATATATATATATATATATAT 7693
 QY 2426 TCACTTCATATATATATATATATATATATATATATATATATATATATATATATAT 2485
 DB 7692 TAAATTAATATATATATATATATATATATATATATATATATATATATATATATAT 7633
 QY 2486 AATCCAGAAATTTCAATTTACGCTGATATGCTGACGCTTCATCCAAACAAATTCAGCA 2545
 DB 7633 ATGTAATATATATATATATATATATATATATATATATATATATATATATATATAT 7586
 QY 2546 GTTCGTGATACATTAATCTCCAAACAGATTAATGATTTTAAATATCTGCTTTCAAT 2605
 DB 7585 TTTCTTTTATATATATATATATATATATATATATATATATATATATATATATATAT 7526
 QY 2606 TCTTCACAGCTTATAGTAATGTAACAGAAATTAATCTTAAACATCATTAATA 2665
 DB 7525 ATAAATTAACAGTATATATATATATATATATATATATATATATATATATATATATATAT 7466
 QY 2666 GGCAGATTAAGATATATATATATATATATATATATATATATATATATATATATATATAT 2725
 DB 7466 AATTAATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 7406
 QY 2726 TTTCAATTAATTAATCAATTTCTTAATATGCTTAAGTTACCTGGCCATTC 2785
 DB 7405 ATTAATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 7346
 QY 2786 TATGATGATGATTAATAGCTTATACATCTTATATATATATATATATATATATATATATAT 2845
 DB 7345 TATTTATATATATATATATATATATATATATATATATATATATATATATATATATATAT 7286
 QY 2846 TTACAGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2905
 DB 7285 TTA-----TTTAAATATATATATATATATATATATATATATATATATATATATATATATAT 7231
 QY 2906 AAAGAGGATCT-TAATTTGAGAGCATTTATATATATATATATATATATATATATATATATATAT 2964
 DB 7230 AATTTATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 7171
 QY 2965 ACCAGCTTTACTTCACTTCAATCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3024
 DB 7170 TATGATATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 7111
 QY 3025 ACTACAGAGATCTTCACTTATTTCCCTATATATATTTTCCCTCGTGAATTAATCTGAGTAA 3084
 DB 7110 TCTAATATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 7051
 QY 3085 TCTTAAT 3091
 DB 7050 TTTAAT 7044

RESULT 20
 AEO15944 300205 bp DNA linear BCT 06-FEB-2003
 DEFINITION Clostridium tetani E88, section 9 of 10 of the complete genome.
 ACCESSION AEO15944 AEO15927
 VERSION AEO15944.1 GI:28204325
 KEYWORDS
 SOURCE Clostridium tetani E88
 ORGANISM Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 Clostridium.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCES
 AUTHORS
 TITLE
 JOURNAL
 FEATURES
 source

1 (bases 1 to 300205)
 Bruggemann, H., Baumer, S., Fricke, W.F., Wierzer, A., Liesegang, H.,
 Decker, I., Herzberg, C., Martinez-Arias, R., Merkl, R., Henne, A., and
 Gottschalk, G.
 The genome sequence of Clostridium tetani, the causative agent of
 tetanus disease
 Proc. Natl. Acad. Sci. U.S.A. 100 (3), 1316-1321 (2003)
 22457253
 12552129
 2 (bases 1 to 300205)
 Bruggemann, H., Baumer, S., Fricke, W.F., Wierzer, A., Liesegang, H.,
 Decker, I., Herzberg, C., Martinez-Arias, R., Merkl, R., Henne, A., and
 Gottschalk, G.
 Direct Submission
 Submitted (31-OCT-2002) Goettingen Genomics Laboratory, Institute
 of Microbiology and Genetics, Georg-August University,
 Grisebachstr. 8, Goettingen 37077, Germany
 Location/Qualifiers
 1. 300205
 /organism="Clostridium tetani E88"
 /mol_type="genomic DNA"
 /strain="Massachusetts"
 /sub_strain="E88"
 /db_xref="taxon:212717"
 complement (162. 683)
 /locus_tag="CTC02285"
 complement (162. 683)
 /locus_tag="CTC02285"
 /notes="tip-3"
 /codon_start=1
 /transl_table=11
 /product="bacterial protein translation initiation factor
 3"
 /protein_id="AA036764.1"
 /db_xref="GI:28204326"
 /translation="NNIIKKGLNMEIEIRKEVRYIAEDGQGVIPTEALKRAEK
 ELDIWIAIAPGKPPVCKIMNYGKPIEDQKKDKKOKVNNVKEIRLSATIEHDI
 GIKANNRKFLEAEDKVKYTVAFRRREHSHSVNGKILKLTLSKVEDCVYERKARL
 ESKNMTVLAERK"
 complement (865. 2784)
 /locus_tag="CTC02286"
 complement (865. 2784)
 /locus_tag="CTC02286"
 /EC_number="6.1.1.3"
 /codon_start=1
 /transl_table=11
 /product="threonyl-tRNA synthetase"
 /protein_id="AA036765.1"
 /db_xref="GI:28204327"
 /translation="MINITLDDGKVIYEKGVKSDIVMKISPALYKXAVGAKINGEI
 AILMEIRKEDSELITLTDDEGRGRTVHTSHLAAQVRLYPAKXALGPAIDNGF
 YDPPIDPFTPEMIEKIEKMAKIVKENLEIRELREBAIKIYKXASPEYKRELI
 EDLPGEVLSFYKQDPSVDCGPHLPSGTGKIKAIKLISVAGATYRGDEKMKLORIT
 ENPWEKRTKAGYDEIRTPVILNELMEIRSGHWDKKNMFTKIDNFPALKPNCPC
 GSILVYKSHLSYKSEFPWRLGLVHRSHSGALHGMVACFQDDAHAFMTLEOI
 KDEILNVILKIDSFYKVFGEYFVLSTRPDSMGSDDEDEVAITNGALNKALEGAOLEY
 KINEDGAFYEGKIDFHLKDCIGRTWOCGTQIDQWDERFELTYVGGDEKRGVLMV
 HRVFGSTIERFTGILIEHFAQFPWLAIPVQKMTITDSQDVANKVINDIKKEGIR
 VEPDRNRKIGIKIRBAQIQKVPYITLIDGKVSBNKVAVSRKSGDGLAISLKEFVA
 KNIYELDNRIYVNSK"
 complement (3117. 4016)
 /locus_tag="CTC02287"
 complement (3117. 4016)
 /locus_tag="CTC02287"
 /codon_start=1
 /transl_table=11
 /product="conserved protein"
 /protein_id="AA036766.1"
 /db_xref="GI:28204328"
 /translation="MKYMAIFKIIYEDDPKEIIEGINALKYFPELEIGVCSIDNNIH
 FWRIVIPSEKLTGKSKNIFHLIANIIVNIYIDKPLKEMTEPIESYFPLNSDKIKE
 IEKIRDLITNNKMKIDSSNIYNNKKNIEIMDEICIKENNEINISGLTFRRKTLK

KEKIIIEKIEVEYIEKEVEDEFTKILKYPVNIQSRIDEVNIITIKKNTIKDNG
 IDTEBELKPDTSNNVSLINDDITLISGLITYPECKTIHCVENCLINKELINTINNV
 ELKAVACNSCTCKREINKOYRK
 gene 4093.4812
 /locus_tag="CTC02288"
 /locus_tag="4812
 /locus_tag="CTC02288"
 /codon_start=1
 /transl_table=11
 /product="conserved protein"
 /protein_id="AAO36767.1"
 /db_xref="GI:28204329"
 /translation="MLGDVVMKYLEVDGRISIEEYVNLMSLENNIKCPISLLYVAI
 SHPDILHLDEKXLIHKDIKEFIKLNLPNIKSLKSLSNYPNMGIGMAN
 LSNVIEHLKTPDPLTENKQKQILNKQGYTCSYAVSDNAILTSQKIAKAE
 VSDVLLIPGDIILPGIDYGFISGCCILKIDMAFFGDLNFTNGEIKELFYKRN
 VPIYKKKGLDRGSLCNE"
 tRNA complement(5057.5131)
 /product="tRNA-Cys"
 /note="codon recognized: UGC"
 complement(5138.5212)
 /product="tRNA-Gly"
 /note="codon recognized: GGC"
 complement(5218.5293)
 /product="tRNA-Phe"
 /note="codon recognized: UUC"
 complement(5403.6491)
 /locus_tag="CTC02290"
 complement(5403.6491)
 /locus_tag="CTC02290"
 /EC_number="2.4.2.21"
 /codon_start=1
 /transl_table=11
 /product="nicotinate-nucleotide-dimethylbenzimidazole
 phosphoribosyltransferase"
 /protein_id="AAO36768.1"
 /db_xref="GI:28204330"
 /translation="WVITMKLLEETLRIRIKADEEVIKRAKMSDLSKPTIGLKKE
 EIAVIGSITGRKVNINKMTVMCSDNIGCEGVACQDELTAALAEVVEITGI
 GVLKRVNADVCVDIGIKSDVKDSRIINKKVAAGTKNMAKEPATREEVKAIETGI
 ELVDFVEGYDLEFTEGAGIGNTAAVSVLSQIDSKIVGSGGLDEGLN
 EAVKQAEINKNKEDVDIVAKIGGPIAGICGPIGAANKRVPIVIDIISAL
 CAYKNNATVDFELISHSABEGIEVYTGALPKCLHMERLQEGSGCLOPFMES
 ALCLMNNMTAELASIVDSERIVDIRK"
 gene complement(6976.7875)
 /locus_tag="CTC02291"
 complement(6976.7875)
 /locus_tag="CTC02291"
 /codon_start=1
 /transl_table=11
 /product="hsp33 (redox regulated chaperone)"
 /protein_id="AAO36769.1"
 /db_xref="GI:28204331"
 /translation="MEGNKADKLKIKATANDQGVRIIGAITTELWNGKIEVHKCSPTG
 AALGRMLTAGLSMGSMLEKEDTIKIDGSGAKVLTAYBEGVAKYIGNPLVH
 LPLNONGLDVGAIGKNGKTVIKDGLDPIYGOVPIYSGEIDDLAYFTVSSOT
 PSAYGLVLYDKDLSIKASGGEIIQMPGADHLLADITIRLEIPISTELISKMSI
 EELIEPIEGMDKILIEGIVPEYTCDSRERKIDALISIGKDLKEIYDEGKTELVC
 QPCNKYIYDEKTEGELIRIKM"
 gene complement(7879.8619)
 /locus_tag="CTC02292"
 complement(7879.8619)
 /locus_tag="CTC02292"
 /EC_number="2.1.1.-"
 /codon_start=1
 /transl_table=11
 /product="methyltransferase"
 /protein_id="AAO36770.1"
 /db_xref="GI:28204332"
 /translation="MECYKQFAQIYDNLIKEDVDYHSNFINDIIVTNHVKQNYLD
 LACGNTTEIISKHFTTAVDLSMDLAEMKLYNGTNGNFIKODISKLELGE
 ENLITCCDSINVLISIEEVSFLRYVNLKDEGIFIPINISYKITEVLGNNTY
 DSEEVFTWENYLNKDIYKYLTFIKKASYSYERPEVAREAYKEMLEKLIKKN

FKIDKNGEYFEKRAVEEKSERIVYLYKK"
 gene 9133.10131
 /locus_tag="CTC02293"
 9133.10131
 /locus_tag="CTC02293"
 /EC_number="1.2.1.11"
 /codon_start=1
 /transl_table=11
 /product="aspartate-semialdehyde dehydrogenase"
 /protein_id="AAO36771.1"
 /db_xref="GI:28204333"
 /translation="MCMKYVALVGTGVMGKPIELERKPEVDNLTLPASASAG
 KIOFCEKREFTVLEKEDNIKIDIALPSAGCISTEYKPIFVRYGATVDNSAW
 RMDPSVLPVPEVPEPDVLSGKITANPNCSTIQAVAKPIYDEYGRITRYSTQV
 VSGAGAGVNDIKDNGVAPKPYALHGNALPHIDSLDNGYKKEKVDNERTKI
 FHEGKITATVRIPEVNSHSESINVELKEFNEDEVELKYNQGITLEDVENVL
 YPMALPASGKDPFVGRIRDRPSINNGINTLVVADNIRKGAALNTIQIAECLIKK"
 gene 10158.11042
 /locus_tag="CTC02294"
 10158.11042
 /locus_tag="CTC02294"
 /EC_number="4.2.1.52"
 /codon_start=1
 /transl_table=11
 /product="dihydrodipicolinate synthase"
 /protein_id="AAO36772.1"
 /db_xref="GI:28204334"
 /translation="MSIFKSGVAIVTPCEGNNPKLEIELIEMHIENSTAYICG
 TTGEASTYREAIKEIKETITVDKVGRIPIVIGTSSNNTRKIEISKAKSIGDGL
 LITPIYNTKQGIIVHEFPAINDVDVPIILYNVGRGNTLPTLVKICDTCSNV
 ALKEASGDLSQIEMKALIKORIDIVSGDDDIPIILSIGIGVSVIANIMPEVHD
 NCELYKGTGKTRALBIQNLTLSTNSLFIETPIPVKIAMNIMGIDVELRLPLCMD
 ESNLMLKKELESHALIK"
 gene 11057.11824
 /locus_tag="CTC02295"
 11057.11824
 /locus_tag="CTC02295"
 /EC_number="1.3.1.26"
 /codon_start=1
 /transl_table=11
 /product="dihydrodipicolinate reductase"
 /protein_id="AAO36773.1"
 /db_xref="GI:28204335"
 /translation="MEVIMVRIIMVGNCKMKGRITITCSKDFNDIEIVAGYDKSTS
 NIDPFVENIHSSVECDVLDLFSRPSLSKSLISYCTKRPVLYCTTGYSKEELNLI

Query Match 3.8%; Score 182; DB 1; Length 300205;
 Best Local Similarity 49.3%; Pred. No. 8.1e-15;
 Matches 606; Conservative 0; Mismatches 600; Indels 24; Gaps 4;

QY 2498 TTCAATTTCGGGTGATGCTGAAGCTTCATTCACAAATTCAGCACTTGGTATACA 2557
 DB |||||
 45599 TTTATTTTATACCTTAAGTTTCTTACCACTTACAGAAATTCAGCGTTTCTCCRA 45658
 QY 2558 TTAACCTCAAGAGATGATGAGTTTAAATACGTCTTCAATCTTCACAGAGT 2617
 DB |||||
 45659 TTTCCATGGAATGTTTATGATATATTCATTAACCTTTTTCATGAGTTTAAAGCT 45718
 QY 2618 TTAAGTATGTACAGAAATTAACCTTACGACTTAAACATCATTAATAGGCAATTAGT 2677
 DB |||||
 45719 TCTCCATATATATATGAGATTAATCAATCTTTCTCATCAATATTTCTGTTTGAAT 45778
 QY 2678 ACTATATCTACTGTAAATCTTTGCTGTAATCATCTCTTAGAGATTTCTAATTA 2737
 DB |||||
 45779 A---TCTCTCAGGAAATCTT---CTTTTATATATATTTCTTCTGTAAGCACTATT 45832
 QY 2738 ATCCATTTATAACATTAAGTATGTTCTTAAAGTTAAGTACCTGCAATCATATGACTGCAAT 2797
 DB |||||
 45833 GCCCTTTCATATATATTTTGTAGTCTCTAATTCATGATATATATTTTCAAGG 45892
 QY 2798 GGTATTAAGCTTCACTTATATACAGAGATCAACCAATTTTATAGATGCA 2857
 DB |||||
 45893 CTATCTAAGCATCCCGCACTTATCTTTTATATCTTTTCTAATTTAATTTAATCTT 45952

QY 2858 TTCTAAATACCTACCAATTCGGGATATCTGTACAAATTTCTACTAAAGCGTACT 2917
 DB 45953 TTTTAAAGATTTCTATTAAATTAAGCAATATCGCTTTTCTCGCCTTAATGGGCAAT 46012
 QY 2918 CTAATTTGGAAGACATTTAATCTATTAATTAATCTCACAAAACCTACAGCTTTTACT 2977
 DB 46013 TCTATATTATTAACATTTATTTCTTAATAACAAATCTGTCTAAATATTTTCTATT 46072
 QY 2978 TCACCTTCAATATCTTTGGAAGAACACAAATTTCTCATCTATTAACAGAAAC 3037
 DB 46073 TCTTCTAATTAATTTATTAATTAAGTGTATTAATTAAGTCTAATTTATTTGAA 46132
 QY 3038 TTACATTTTCCATTAATTTTCCCTCCGTAATATCGAATTTCTAATTTGGA 3097
 DB 46133 CTAGACCTATAGCTTTATCTCTTATCTGTAAATCTTTAAAGTTAGTTGAA 46192
 QY 3098 TCGATACGTAATCTGTACTTCACTAATTAATTAAGTACCAATGATTGCTCTATA 3157
 DB 46193 TTTATGTCATATCCCTAATTTCACTAAACAAATTTCCCTTTGAGATGTTCAAT 46252
 QY 3158 ATTCCAAATAGAGAGTACTCTATGAGCAATATATTTACTTTCCCACTAACAAAT 3217
 DB 46253 ATACCTTTTATCTCTACGCTCCAG-----TAAAGCTCCTTTTATGACCAAT 46306
 QY 3218 ATATTAACCAATATTAATTAAGTGTAGACATATAGATACATATTAATTAAGAGTA 3277
 DB 46307 AATTCACCTTTCTAACAATTTCTGGAATGCTGCAACATTTATAGCATTAAGAGACCT 46366
 QY 3278 TCATACCCCTTGATTTTATTTATTAAGCTAGCACTACTCTCTTCCAAACCTGGC 3337
 DB 46367 TCTTTTCAATTTCAATTAATTAATTAAGCTAGCACTACTCTCTTCCGTAACCTT 46426
 QY 3338 GAACAGTAATAGATATACGACTCGATGTAGTACTCTTTATTAATCATATCTTCCAA 3397
 DB 46427 TCTCAGTTAATAGATACATTAACATTAATTCCTTCACTCATATTAAGTCAAAATC 46486
 QY 3398 TTACGTATTAACAGGAGCTTACCGCATTTTCT-----AATCTCAATGCT 3445
 DB 46487 TTGTTCATTTCTTTGAAATTTCCATTAATTAATTAATTAAGTCAATTTGG 46546
 QY 3446 GATTCACATCATATTTCTCTAGTATCTACAGCTATAGCTCTCTTACACT 3505
 DB 46547 CTGCTTAATATTTATTTGGAATTTAGATTTATTAATTTATTTGCTTGAAT 46606
 QY 3506 AACTTAATCTCTCTGTAAGAGCTTTTCTAATTAATTAAGACCAATATGCA 3565
 DB 46607 AAACCTAATCTCTCTAATTAATTTGTTAGTTACATTAATTAAGACCAATTTCTATA 46666
 QY 3566 GACTTACAGAGTGGCAATATTAACATGCCCACATTAATTAATTAAGCAATTAAGA 3625
 DB 46667 GCTTATTAAGAAATTTCTATTTGCTTCACTAGCTCTAATTAATTAATTTGCTTT 46726
 QY 3626 TACCTTTCTTAAGCTTTTCCAGTACACTTAATCAATTAATTAATTTACCAT 3685
 DB 46727 TTTATTTCTTATTTCTTATTTCTTATTTCCAAACCAATCTCTGACCAATTTAAATCC 46786
 QY 3686 AATTCATTAATTAAGCAATGAGCTTTT 3715
 DB 46787 AATTAATTAATTAATTTCTTATTT 46816

RESULT 21
 AC134919/c 145992 bp DNA linear PRI 07-AUG-2003
 LOCUS DEFINITION Homo sapiens BAC clone Rpl3-485B10 from 4, complete sequence.
 ACCESSION AC134919
 VERSION AC134919.3 GI:33386922
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 145992)
 Sulston, J.E. and Wilson, R.
 Toward a complete human genome sequence
 Genome Res. 8 (11), 1097-1108 (1998)

REFERENCE

2 (bases 1 to 145992)
 McMillan, M., Haglund, K., Du, H., and Cotton, M.
 The sequence of Homo sapiens BAC clone Rpl3-485B10
 Unpublished (2001)

REFERENCE

3 (bases 1 to 145992)
 Waterston, R.H.
 Direct Submission
 Submitted (02-OCT-2002) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA

REFERENCE

4 (bases 1 to 145992)
 Waterston, R.H.
 Direct Submission
 Submitted (20-NOV-2002) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA

REFERENCE

5 (bases 1 to 145992)
 Wilson, R.K.
 Direct Submission
 Submitted (01-AUG-2003) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA

REFERENCE

6 (bases 1 to 145992)
 Wilson, R.K.
 Direct Submission
 Submitted (07-AUG-2003) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Aug 1, 2003 this sequence version replaced g1:2510186.

REFERENCE

----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu>
 Contact: sapiens@watson.wustl.edu
 Summary Statistics
 Center project name: H_FH0485B10

COMMENT

NOTICE: This sequence may not represent the entire insert of this
 clone. It may be shorter because we only sequence overlapping
 clone sections once, or longer because we provide a small overlap
 between neighboring data submissions.

COMMENT

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate
 chemistry, or covered by high quality data (i.e., phred quality >=

COMMENT

30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by sequence
 from more than one subclone; and the assembly was confirmed by
 restriction digest.

COMMENT

MAPPING INFORMATION:
 Mapping information for this clone was provided by Dr. Wes Warren,
 Department of Genetics, Washington University, St. Louis MO. For
 additional information about the map position of this sequence, see
<http://genome.wustl.edu>

COMMENT

SOURCE INFORMATION:
 The RPL3-13 Human Female BAC Library was constructed using improved
 cloning techniques developed by Kazuo Osoegawa. The library was
 generated by Baochun Zhao in our laboratory. Construction was
 funded by a grant from the National Human Genome Research Institute
 (NHGRI, NIH) (#R01HG01165-03). The library was generated according
 to the new NHGRI/DOE 'Guidance on Human Subjects in Large-Scale DNA
 Sequencing'.

COMMENT

Female blood was obtained via a double-blind selection protocol.
 Female blood DNA was isolated from one randomly chosen donor (out
 of 10 female donors) and partially digested with a combination of

[illegible]

[illegible][illegible]

```
gene /sub_species="nucleatum"  
/db_xref="ATCC:25586"  
/db_xref="taxon:190304"  
CDS 116..859  
/gene="FN0892"  
116..859  
/gene="FN0892"  
/EC_number="3.1.3.3"  
/codon_start=1  
/transl_table=1  
/product="phosphoserine phosphatase"  
/protein_id="AAL95088.1"  
/db_xref="GI:19714455"  
/translation="MKIKMIAAFDIDGTYRNALLIEHFKLVKVELPDIDIOYRLV  
EAVYIMPTRGDYDDYLLDTOLYYVAKLPYKANDPISNOYLLKGRVYTYRE  
MTBWKKGHHVFTSGSPFLVSRMAKKGVDGCVIEIDETQTSFGKLTQPMW  
DSAKQEAIEHFIKKYNIDLSKYAGDTNGDFMSLVSNPAINPSKELITRVND  
ENLSKTIQIIIEKRVIVKLSNDELIEF"  
gene 894..1316  
/gene="FN0893"  
894..1316  
/gene="FN0893"  
/codon_start=1  
/transl_table=1  
/product="Hypothetical protein"  
/protein_id="AAL95089.1"  
/db_xref="GI:19714456"  
/translation="MKIKNEVVALQIIYYTLNRDQKIISSNEISAENIPLPCIR  
IILKEKAGVIERGAKGVYLRDRKRLFRPIIETIIDDYVLQPCIDSSITCSTR  
GADCSIRALKKIQDDLDPDKINFIDLVENNASLQI"  
gene 1347..1883  
/gene="FN0894"  
1347..1883  
/gene="FN0894"  
/codon_start=1  
/transl_table=1  
/product="unknown"  
/protein_id="AAL95090.1"  
/db_xref="GI:19714457"  
/translation="MDIQENKYYADDMGKEVYVKNLIANQAGIILYLIPLFLVK  
NGILKGDGTFRLIVCTIFLFLNPLSQKLPFLKTKAKSIHNDQSYPTLVQFQV  
IFMOGKPSVELDSKIYKIVLYKYSVLMETNSNGIMIKYDPTKCNFDPFBEIKE  
NCKKAKIIVKRSYIFGL"  
gene complement(1929..2393)  
/gene="FN0895"  
complement(1929..2393)  
/gene="FN0895"  
/codon_start=1  
/transl_table=1  
/product="unknown"  
/protein_id="AAL95091.1"  
/db_xref="GI:19714458"  
/translation="MDTKLKSKITDSFPKGIKTIENTYKNHDEKYSGCRIQEGYD  
DYLKIVFKRKIEYKDDPFWSDTPDFEIIFOELNEVKRDPVEEIVPEIKLSFBEEN  
DILDKTKEFFIENNERKKELSKMENIYKEVIFDEKAMDEENKLLI"  
gene complement(2417..2800)  
/gene="FN0896"  
complement(2417..2800)  
/gene="FN0896"  
/codon_start=1  
/transl_table=1  
/product="Hypothetical protein"  
/protein_id="AAL95092.1"  
/db_xref="GI:19714459"  
/translation="MKKEKVVVEVREGVVSEEDVLIAGLIISFLIGSPFTIIPKOV  
FLAVNKEFFIATLTKMGDDPKIKIHYSIDENVQMKKGLLNGCLILTHSDGGRB  
NYRAMKPAFGSLAANNFKALQKFG"  
gene 2903..3571  
/gene="FN0897"  
2903..3571  
/gene="FN0897"  
/codon_start=1  
/transl_table=1
```

```
gene /product="Hypothetical cytosolic protein"  
/protein_id="AAL95093.1"  
/db_xref="GI:19714460"  
/translation="WRLEKDFPKNKIDFKKLEFGFKLIDNSYYHTLSLKNQFRMS  
VKINLDSIFTEIIDETNEPYLILMLKSGSYSEKYYKAYENAVLEITQKCEPEDI  
FKANYTKRIIDVYKKGDELPELMKSPKNAVIRKSSKNQYAVITLISRKIGLDS  
DKIIEVINLHNSSEIYKLIIDCKKIFPAYHNNKTKWCSICDGYLELEIYKLIIDISY  
ELAK"  
gene 3646..5517  
/gene="FN0898"  
3646..5517  
/gene="FN0898"  
/EC_number="4.1.99.-"  
/note="spore photoproduct"  
/codon_start=1  
/transl_table=1  
/product="Hypothetical protein"  
/protein_id="AAL95094.1"  
/db_xref="GI:19714461"  
/translation="WLYIVTALYTEAKPLISLNLKKNQTYTKRQVSENNIKLISG  
TKIKSATATLYLSENKIKNDYITNGFIASSNNSQLGIYYSKIQNAYSPTR  
YPMYKXNFBLSLITPDKIIEKTEVETIDENAYGFOTASIFPKDKIPLKIV  
SDILKENVEDRLIDFKODNLNKS YKIVPLKFINI PNNKNENNNQDLTKV  
LENKLSLMTYEFENILKYLKIKYGNIDILKKEENIEVNSKVGKKIIEEIKESKL  
NNKVEIERKSPNNKNSLNNRFSHIVYEKTLNNKTLIELSKFQKVIIEIDNKE  
VSSNNQDFHLQKLGKLIASNNKPMIYEGAVVCESENNFYTSIIINCVDCEY  
CYLGIVSSGNIVIFVDIEKVEEELYSKLTLYLCVSYDTDLAIENCGSPFKW  
VHFIEDKVKYKTEILTKSGNIDKFLNKLPLNFIIFALISEBNALAKNEXYASFRN  
VATIEIDQKMKVATCIDPLISNFKANSQMEIYFNSIDEKVIIDISIGVRIS  
KEYLKNRQNNQNSILYFPPECIDGYTYSBKTSWIMNIFKQFLKYNINKLYT"  
gene 5527..6243  
/gene="FN0899"  
5527..6243  
/gene="FN0899"  
/EC_number="1.1.1.100"  
/codon_start=1  
/transl_table=1  
/product="3-oxoacyl-[acyl-carrier protein] reductase"  
/protein_id="AAL95095.1"  
/db_xref="GI:19714462"  
/translation="MKIALVGTSGTSGIYEISKRLKNQYTVYIGRNFKNENIFK  
EYENFIVTQDLSKLDLEKTLHSKKIKQDLVNSAGIYFGGHEENVSKIKMT  
VNLQAPVYISQFPLRTKENKGIITNSVANNKSPASVYSATKAGLSQPSKLE  
EVRNDVYITITPMTKTNPYANNITPECDDEKAYIKMDIGNTIFILNQSNIY  
FDITIKPQREKIKVKKKE"  
gene 6250..7032  
/gene="FN0900"  
6250..7032  
/gene="FN0900"  
/EC_number="3.-.-.-"  
/codon_start=1  
/transl_table=1  
/product="Metal dependent hydrolase"  
/protein_id="AAL95096.1"  
/db_xref="GI:19714463"  
/translation="MKISILGSGAGNSTFVEIDYKLVDTGSCCKTEELKRIKQ  
KLSDISALITHESHDHNGAVLARKNDITITPEBSYAGAVLQGDIDSLNFIID  
GDFIINDKVSYPDMWDAPRTGFKELQNLNKRIASDTIGYTNIVREYFQVDA  
MYTESNDPNTLACAYPMNKERYKSGNGLSNNECKPTEKMTDTLKKYFLAHVS  
KDSNNISLIRKETLEDEFQMIKRPVCEITTDNVTKLPDIDE"  
gene 7042..7629  
/gene="FN0901"  
7042..7629  
/gene="FN0901"  
/codon_start=1  
/transl_table=1  
/product="DNA polymerase, bacteriophage -type"  
/protein_id="AAL95097.1"  
/db_xref="GI:19714464"  
/translation="WDEISLWEDLKFPAGSIGNELJPKDQOEYVIGGDMNADILFV  
GNDPLVLAEDVYKSSGSAFLIRLDDVYVLEBEYVITLTSKREIKIKNFNEERK  
KLIDLPQOILLISPKIVVFLGKVAQALENKEIDPDERGQPFKMRGDIETVLTGYV  
ETVIAKRNDSKKAAIALNPLDMKNIERLNND"
```

gene
7619..8164
/gene="FN0902"
CDS
7619..8164
/gene="FN0902"
/EC_number="6.3.3.2"
/codon_start=1
/transl_table=11
/product="5-formyltetrahydrofolate cyclo-ligase"
/protein_id="AAJ95098.1"
/db_xref="GI:19714465"
/translation="MNKKKARLLIKERRMLSKRYIDVASDXIEKLONDEPKNK
TWMSYMDPKNEKTRINTPIKNSKIIIVLEKVDPEKEMVITEKNQYIYSPENKED
DDEKRGSDIVITTPGAFDRKKNVGRGRIYDFVQPNARKIAIAFEKQIDEG
IETDKYDKVDLITEDIK"
8186..9157
/gene="FN0903"
8186..9157
/gene="FN0903"
/db_xref="GI:19714465"
/gene="FN0903"

Query Match 3.7%; Score 175.8; DB 1; Length 12277;
Best Local Similarity 50.0%; Pred. No. 1.5e-13;
Matches 527; Conservative 0; Mismatches 517; Indels 11; Gaps 3;

QY 627 AAAAAATGAAATCGATTATTTAGTGCAGATCATTTGGCAGCAATAGCAATTGAC 686
DB 12153 AATGTGTAATATATCAGTTATTTGTTGCGAGATGGGAAATAGCAATTAGCTATTTAT 12094
QY 687 TGTGACGACATGATATATCATTTAATGAGGAGTATCATGAAATATTAACATA 746
DB 12093 TACATTAATATGACATTAATTAATATGCTTTTGAATTAAGAGAGCTGAGAT 12034
QY 747 TAAACCTTACCGAAATTTAAATATTTTACCAATATCATCTACCAACAACAT 806
DB 12033 TAAATTAATATGACATTAATTAATATGCTTTTGAATTAAGAGAGCTGAGAT 11974
QY 807 ATGCAACGACATATATGAGAGATTTATCTGACACATATCATGATTTATTTAATA 866
DB 11973 AAGTACGATATATTTAAAGAGAGCT--TGATTAATTAAGATTTTGGTGGTGAG 11917
QY 867 TTCTTACACATTTGCGACCATATGTCACAAATACACACAAAGCATATGTGTA 926
DB 11916 TTCTTACAAACATTAAGATGTCATTAATTTTAAAGATTTATCAAGATATATC 11857
QY 927 AAAAAATCAGATTTATTTAGTAAAGATGCAATTAATCATCTCAATTTCCCA 986
DB 11856 AATTAATTTGTTGTAAGAGAGCTTAAAGAGATATCAATTAATTAAGATTAAG 11797
QY 987 GTGAAATAGCAGAGAAATTTTACATATATCAATTTTATCTCTGTCAGATT 1046
DB 11796 TTAATGAGAGAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 11737
QY 1047 TTGCTTAAAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1106
DB 11736 ATGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 11677
QY 1107 AACTGTGGAATCATGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1166
DB 11676 AACTTACATTAATTTTACAAATTTTATTTATGATGATGATGATGATGATGAT 11617
QY 1167 AAGATATTTAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1226
DB 11616 CTGATATGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 11557
QY 1227 GAATATTCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1286
DB 11556 GAATAGCTATGCTGCTAAATTTAGAGAGATTAACAAAGAGAGAGAGAGAGAG 11497
QY 1287 TGAATGAATTAATTAACATATATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1346
DB 11496 TAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 11443
QY 1347 TTGCTCATCATGCTTTGAGATCTAATTAATTAACATGTAACAGAGAGAGAGAG 1406

DB 11442 ATGACATACTGTTAGAGAGATTAATGCTTATCATGTCAGATATGACAGATAGAAATA 11383
QY 1407 TGGCTTTTGACATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1466
DB 11382 GAAAGAGAGATCTTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 11323
QY 1467 TAAAGCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1526
DB 11322 ATATGAGAGAGAGAGATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 11263
QY 1527 ATGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1586
DB 11262 ATGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 11203
QY 1587 ATAAAGCAGATATCAATATTAATGCTATCATCTTGAATATTTCAACAGATG 1646
DB 11202 CAGAGAG--TGTAATGAATTAATGATTAAGAGATTAAGAGAGAGAGAGAGAGAGAG 11145
QY 1647 ATAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1681
DB 11144 GTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 11110

RESULT 24
AP004827/c 333750 bp DNA linear BCT 24-APR-2003
LOCUS Staphylococcus aureus subsp. aureus MW2 DNA, complete genome,
DEFINITION Staphylococcus aureus subsp. aureus MW2, section 6/10.
ACCESSION AP004827 EMBL000033
VERSION AP004827.1 GI:21204509
KEYWORDS

SOURCE
ORGANISM Staphylococcus aureus subsp. aureus MW2
Staphylococcus aureus subsp. aureus MW2
Bacteria; Firmicutes; Bacillales; Staphylococcaceae.

REFERENCE
1 Baba, T., Takeuchi, F., Kuroda, M., Yuzawa, H., Aoki, K., Oguchi, A., Nagai, Y., Iwano, N., Asano, K., Naito, T., Kuroda, H., Cui, L., Yamamoto, K. and Hiramatsu, K.
Genome and virulence determinants of high virulence community-acquired MRSA
Lancet 359 (9320), 1819-1827 (2002)

TITLE
JOURNAL MEDLINE
PUBMED
2 (bases 1 to 333750)
12044378
Director-General, Biotechnology Center, Aoki, K., Oguchi, A., Nagai, Y., Asano, K., Iwano, N., Baba, T., Kuroda, M., Hiramatsu, K. and Kikuchi, H.
Direct Submission
Submitted (06-MAR-2002) Director-General, Biotechnology Center, National Institute of Technology and Evaluation, Biotechnology Center, 2Chome 49-10 Nishihara, Shibuya-Ku, Tokyo 151-0066, Japan
E-mail: biocentre-go.jp, URL: http://www.btc.mite.go.jp/, Tel: 81-3-3481-1933, Fax: 81-3-3481-8424

FEATURES
source
1..333750
/organism="Staphylococcus aureus subsp. aureus MW2"
/mol_type="genomic DNA"
/strain="MW2"
/db_xref="taxon:196620"
complement(278..619)
/gene="MW1341"
complement(278..619)
/gene="MW1341"
/note="ORFID:MW1341"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAB95206.1"
/db_xref="GI:21204510"
/translation="MNITKREVDNLEQSDIVPAKNTLGSSEAKPYLDSEYKIIDY
FRVNDITSDIDKLNQIPVPMNFKERYDYMIERKHIFMGYRQMKTKTELLIGNA
SYOTRLKNQV"
complement(624..1283)

gene
11442 ATGACATACTGTTAGAGAGATTAATGCTTATCATGTCAGATATGACAGATAGAAATA 11383
QY 1407 TGGCTTTTGACATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1466
DB 11382 GAAAGAGAGATCTTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 11323
QY 1467 TAAAGCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1526
DB 11322 ATATGAGAGAGAGATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 11263
QY 1527 ATGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1586
DB 11262 ATGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 11203
QY 1587 ATAAAGCAGATATCAATATTAATGCTATCATCTTGAATATTTCAACAGATG 1646
DB 11202 CAGAGAG--TGTAATGAATTAATGATTAAGAGATTAAGAGAGAGAGAGAGAGAGAG 11145
QY 1647 ATAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1681
DB 11144 GTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 11110


```

CDS
    /gene="nch"
    complement(624..1283)
    /gene="nch"
    /note="ORFID:MW1342"
    /codon_start=1
    /transl_table=11
    /product="endonuclease-like protein"
    /protein_id="BAB95207.1"
    /db_xref="GI:21204511"
    /translation="MWSKKALMIDVIANMPDAECGLKHNPELITAVLSAOCCT
    GELPQHKLESISLAGKRTANVWSVAFDESLVDYHVERVSRGLINKKKNVRQ
    VEDRCSVIPRDRMNRSHQLFFGRHOLAKPKDCIPLEDCEGKRYKASLKE
    A"
gene
    complement(1273..1959)
    /gene="MW1343"
    complement(1273..1959)
    /gene="MW1343"
    /note="ORFID:MW1343"
    hypothetical protein, similar to chromosome replication
    initiation protein dnaD"
    /codon_start=1
    /transl_table=11
    /protein_id="BAB95208.1"
    /db_xref="GI:21204512"
    /translation="MDKYOQKARPVVIRRELLDHSYDGLDEBDIVILKTIYSETS
    NKPSIELQKSTQPDITWVIONLQRELELQVQDEGRFTEYMLDPPFEKL
    SHLQKQSMETEQNSKEKFKQLFRVLEDTPARPLSPYEIKTNQWIDVDGDAIIO
    AALDEANSLNKLSPFYKMDRILMMKKNNVKIIDSRIKREKPKMTHVTKVYKPD
    MNGENLDKRG"
gene
    complement(2287..3579)
    /gene="aasn5"
    complement(2287..3579)
    /gene="aasn5"
    complement(2287..3579)
    /gene="aasn5"
    /note="ORFID:MW1344"
    /codon_start=1
    /transl_table=11
    /product="asparaginyl-tRNA synthetase"
    /protein_id="BAB95209.1"
    /db_xref="GI:21204513"
    /translation="WKTITKQKDHLDNDVTIGAWLNKRSKGIAPLQLRDGTGFGQ
    GVVAEVEDEEVLAKETIQESSLYVTGITTEDRSDLGYEQVKSIEVISEADYP
    ITPKNGTEFLMDRHLMLRSKQKAVAKIRNEVIRATVEFNKDFGTFPLITLS
    AEGTSELPHTKYFDQDAFLSOSQGLYLAAMAAKVFSGFPRAEKSTRRHLIE
    FMWIGEMAFTHAESELEIOEGYTHVAVKLEKTLIERDTSKLEKVAAPFR
    ISYDAIRIKSEGGPDIEKMGDPCGAPHEGTLAHYDLPVETIYVPTKIKRFPVQBP
    ENBEYVLCADLIAPRGYGTITGSESRVDLELLEORVKEHGLDEBAISYIUDLRKYS
    VPHGFGGLGEISTVAMISGVEHRETAFFPRLNQLYF"
    complement(3901..6594)
    /gene="ding"
    complement(3901..6594)
    /gene="ding"
    /note="ORFID:MW1345"
    /codon_start=1
    /transl_table=11
    /product="probable ATP-dependent DNA helicase ding"
    /protein_id="BAB95210.1"
    /db_xref="GI:21204514"
    /translation="WGMATYAVVDLETTGNQDLPDDIIQIGITFVRNNQIITFWSMI
    RNLIEIPFIQALSTIEENMLQAPYENQVACIYDKIKDCIPIAHVNDQIDITFWSMI
    RDNIOYRPKVITDLEIFKIAPTDYSYOLSELARHGHTLANARABEDATTK
    LMITLAFKPEKPLDITLKQLYYSKQLYDYDFFEMVROYDAKPLDKSEKBEQIT
    IRQYDFKPKPTNNNGSLKSLYSKAVDQGLYTPQQYLAETILLDQLMSEKMTA
    SLGSGSLAYLLAALMNIETGKVMISTNKLQSQLEKDI PANMEALNFKINALI
    IKSQSDYSLGSLQKLDPTSNYENVTIKQMLIWIETPSGDIQELNKGQKQMF
    DQIETVYPARHDVYVNIKRANQIQIGITNHAHLHSDVENSIOLEPDCIVDA
    HLLPYALNOVTNELSVADIKYQGLIGKNENKLLKALIDOLEKORITLKKLDAPID
    FLKASMEIHEINQLESTIETIINDVDDIHRPHVFTRETDILKDDAAID
    KLKLTLEIENGISHTYVSLKQKLLYLDKRYKATFQSKAGHTSIFISKLSQSGSTR
    IYVDYAKVDVTKQVDEKFKSLIFISGLTKFNHSEFAKQLENKDVHNTFENSTIR
    QSAKTSVFIQSDVASVQYKNIDEVASIVYILLIETITSSKCLVLFYSKMMHWQ
    DMNLPEFEDVYVLTQQQNONYKIVQGFNNFDAILLGSTFPEGDFDPANGIKCMV
    IAKLPFMKHNNAKWLMDSEFTSEKFEYVLPDAVTRFQGIIRLIRESENGRGLIVSD

```

```

gene
    complement(6618..7589)
    /gene="MW1346"
    complement(6618..7589)
    /gene="MW1346"
    /note="ORFID:MW1346"
    /codon_start=1
    /transl_table=11
    /product="truncated hypothetical protein"
    /protein_id="BAB95211.1"
    /db_xref="GI:21204515"
    /translation="MSKYSQVLYQLLYKKNPVYISQGSIAESLNTSRIVKRYTQQL
    LBGCKIDSVNHKHLQQLPDITWYQGIIDQYKSSALPFPSSVYDSISTQLAASKL
    VGNOSFFPILSDQIKRGFRFNRHWSSSGQGLIMSVVIRPVAFSMISKPMIPLALG
    IRDAIOHFSQDEYKVPNDIYIDNGKVGCFIEMVANNNGIEALICGIGINTQOLE
    NPDESIRHRTASVQLHDKKIDRYOFLERLOIEKRYQFPLTPSEIRREYIASN
    IWRITLLFTENKQFQGAIDIDYDGYLVROBAGSHRLISADIDF"
    complement(7576..8778)
    /gene="MW1347"
    complement(7576..8778)
    /gene="MW1347"
    /note="ORFID:MW1347"
    hypothetical protein, similar to poly(A) polymerase"
    /codon_start=1
    /transl_table=11
    /protein_id="BAB95212.1"
    /db_xref="GI:21204516"
    /translation="MDKSLFEQARPLEQIDNGFEAYVYGGSVADYVWGRNHDIDI
    TTSATPDEIESIFSHITPVGKEHGTINVFNDENYVTFPRAEDYVDHRRSGVTFV
    RDLVEDLORDPDAIMADTAVKYDYDQGDQNNRIIRFVGAEEFQDADLAMI
    RCLRFOSQSPDIATETEPAMRIOMADIKESLIEIYELTKMGINVERKFNELKS
    IKAFTNMPYFEHLDNQNIWTAIDILHILAIYSVKPDINYSIKPLKISNROVQINO
    YIOINNALPSITTEQLKRFVYDIDHLIKRWAVADYKANDIGCHEPLIYNLOTID
    ETHLRPLMNRKDMNVNGVLAHLNANSGLKQVLQIETAIYTVGKSNSETILK
    WVDNHYKI"
    complement(8783..9925)
    /gene="MW1348"
    complement(8783..9925)
    /gene="MW1348"
    /note="ORFID:MW1348"
    hypothetical protein, similar to lipopolysaccharide
    biosynthesis-related pr homolog"
    /codon_start=1
    /transl_table=11
    /protein_id="BAB95213.1"
    /db_xref="GI:21204517"
    /translation="WKTIGTCYPSMGSGSITITELGKLAENGHEVHTSNIPIRIR
    KLPNNITFQVAVNOYAVQYPPYDITLSTKIAVYKEVGLDLMLMHTAVPAIIGIL
    AEMSGKDIKIMTTLHGTIDIVLYGDHSLQGIKIKGIESDITVTSKSLAQBTEHII
    ENKKEIIPYVRENREPEFTKNTALKSQFGIAPDEKYLIVHSNPROVSRIDITILEF
    AKVREKIPSKLILISDGPFLVPMRQUTLELANEEDVLFLGKQDCVSEFQSLDVLILL
    SEKVSFGTLTLEAMKTGVVIGSNAGIKREVIKHSEGTGVVDVQCDSSADYAIRLLE
    DVLVYKLOXKMLADIABERFSGELITDQYEVYYQKMLNHNKSKGE"
    complement(10167..10484)
    /gene="MW1349"
    complement(10167..10484)
    /gene="MW1349"
    /note="ORFID:MW1349"
    /codon_start=1
    /transl_table=11
    /product="conserved hypothetical protein"
    /protein_id="BAB95214.1"
    /db_xref="GI:21204518"
    /translation="WKSVMQREVEDEYIGQFKTGFSPLANLALTEVEGLAREIN
    HLYGEKKKDSSEANTIRAEIGDNLVILLCLANSWIGIMTESFNETMEKFTYRDRNF
    ERK"
    complement(10820..11518)
    /gene="MW1350"
    complement(10820..11518)
    /gene="MW1350"
    /note="ORFID:MW1350"
    /codon_start=1
    /transl_table=11

```

/product="conserved hypothetical protein"
 /protein_id="BAB95215.1"
 /db_xref="GI:21204519"
 /translation="MSIFMTIYFVILMLVPLMHOKHVKSNYKSVQVRSQKTR
 EVALLIANGIYDVEVKGSEFLDHPKRVVSLSPAYDPSVAGTIAHNG
 HAIDHOGWIFRPAALVPVANNLSSISYIMIGITLITLIGAFSGSTALMIGLIM

Query Match 3.7%; Score 174.4; DB 1; Length 333750;

Best Local Similarity 50.9%; Pred. No. 7.1e-14;

Matches 533; Conservative 0; Mismatches 491; Indels 24; Gaps 4;

```

QY 628 AAAAAATGAAATCAGATTTTGGGCGAGGATCATTGGCCAGCAATAGCAATTCGCT 687
DB 23025 AATGACTAAATATCCGTTTGGTATGGAGATTGGACAGCCCTTGCAATGTTCT 22966
QY 688 GTGACGACATGTATATCATGTTAATTGAGGAGCGTATCATGAAATATTAACATAT 747
DB 22965 TCGAGAAATGACATGATGTTTGTATGGGGTAAATCAAGATGCTGTGATGAT 22906
QY 748 AATCACTTACCAAAAAATTTAAATTTTACCCCATATCATCTACCAACAACATATA 807
DB 22905 AATATCATGTCTATCAATAAAAATTTAAATATCCGAATTAAGATGTTAATCATCAT 22846
QY 808 TCGAACCCAGCAATATGACGAGATTTATCTGACACAACATCATSTATTTCTTAACAT 867
DB 22845 CGCTACTTACATATGACCAAGCAATTCATTTGACATATTTACTTAATG---CTT 22789
QY 868 TCTTACACAACATTAACGACCATATGACACAATTAACAACAACGATATGTATA 927
DB 22788 ACCTACTTAAAGCAAGCGAAGATGCTTCCAAATTAATATGATGAGC-----TGACCTC 22735
QY 928 AATATCTCAATATTAATTTAGTAAAGATTCGAAATTAATCATCATCAATTTCCAG 987
DB 22734 TAAAGAACTTTTATATCATGTTGCTAAAGATTTAAATGAGCGTTAAACGTGCTC 22675
QY 988 TGAATATGACAGAAATTTT-----ACATATTAATCCAAATTTTATATCTCTCG 1038
DB 22674 AGAAATGATGAAAGATTCATTTCACTGAATATATGACAGATATGCGGTGTCAGG 22615
QY 1039 TCCAAATTTTGTATAGAAATTTGACAGACATCTTCTGTATAGTACTGCTGGA 1098
DB 22614 GCCAAGTCAATGCGAAGAGTGTATGACAGACCACTACAGTGTGCTTATCAAAA 22555
QY 1099 TAATTAAGAACTTGTGATCATCTATGATGAAACATATGATGATGTTCTTAAATAT 1158
DB 22554 AGATTAAGATGTAAGTAATTAAGCAAGATTTATTAATGATGATTTTGGCTGTA 22495
QY 1159 ATACCATCAAGTATTTATGATGATGATGATGATGATGATGATGATGATGAT 1218
DB 22494 CACGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 22435
QY 1219 TGCATGTGGAATATCGCTGGAATAATTTAGTAAATATGCTGTGCTATGTTATAC 1278
DB 22434 AGCAAGTGTATCGTATGATGATGATGATGATGATGATGATGATGATGATGAT 22375
QY 1279 TAAAGCATGAATGAATTAATAAATCTATATATGCAAAAAATCATCATGATCTTCA 1338
DB 22374 TCGTGTGTTAGCGGAATTAAGTA-----TTAGTGAAGAAAGTAGTGCCGCTCAT 22321
QY 1339 TACATTAATGTGTCATCATGCTTGGAGATCTATATTAACATGTCACAGAAACCTC 1398
DB 22320 GACATTTTCTAGTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 22261
QY 1399 AGCAATATGCTTTTGAATGAATATGAAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1458
DB 22260 TCGAATATTCATAGATATGAATTAACCTTGACAAAGTGAATCAATGATCAAGATATC 22201
QY 1459 TCACAACTTAAGCTTTTGAAGAGACAGATCTGTAACCACTGATATCTATGACAAA 1518
DB 22200 TGAATTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 22141
QY 1519 AATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1578

```

```

DB 22140 AGAAAAATGTCATATGCCAATTAACAATGATATATATAGATATTTGAAAAATAT 22081
QY 1579 ATCACTAGATTAAGCCATATCAACAATTAATTTATGTCATATATCTTTGATTTCA 1638
DB 22080 CTGATGAAGAAATGCTTAAGATTTAATGAGCGCGATTAATAATCTGATTAATATG 22021
QY 1639 ACAATATGATTAATTAACCGTCAAAAT 1666
DB 22020 GATTTATTAACATATATCGCTAATAT 21993

```

RESULT 25

AC005504 104992 bp DNA linear HTG 01-APR-1999
 Plasmodium falciparum chromosome 12, *** SEQUENCING IN PROGRESS
 *** 3 unordered pieces.

AC005504 3 GI:4558584

HTG; HTGS PHASE1.
 Plasmodium falciparum (malaria parasite P. falciparum)

Plasmodium falciparum
 Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.

1 (bases 1 to 104992)
 Hyman,R.W., Fung,E.L., Qin,F., Tanaki,T., Kurdi,O.B., Conway,A.B.

Hyman,R.W., Fung,E.L., Qin,F., Tanaki,T., Kurdi,O.B., Conway,A.B.
 Plasmodium falciparum 3D7 chromosome 12

Unpublished
 2 (bases 1 to 104992)

Hyman,R.W., Fung,E.L., Qin,F., Tanaki,T., Kurdi,O.B., Conway,A.B.
 Direct Submission

Submitted (21-AUG-1998) Stanford DNA Sequencing and Technology
 Center, Stanford University, 855 California Avenue, Palo Alto, CA

COMMENT

On Apr 2, 1999 this sequence version replaced gi:4337172.
 NOTE: This is a 'working draft' sequence. It currently
 * consists of 3 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 58642: contig of 58642 bp in length
 * 58643 58842: gap of unknown length
 * 58843 91011: contig of 32169 bp in length
 * 91012 91212: gap of unknown length
 * 91212 104992: contig of 13781 bp in length.

FEATURES

1. 104992
 Location/Qualifiers
 /organism="Plasmodium falciparum"
 /mol_type="genomic DNA"
 /db_xref="taxon:5833"
 /chromosome="12"

ORIGIN

Query Match 3.7%; Score 173.8; DB 2; Length 104992;

Best Local Similarity 46.7%; Pred. No. 1.3e-13;
 Matches 1029; Conservative 0; Mismatches 1122; Indels 52; Gaps 13;

```

QY 345 AATATATAGAGTCTTATTTGACATATCAATTAATCAATCAATTAATATAT 404
DB 72744 AATACATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 72803
QY 405 ACTATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 464
DB 72804 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 72863
QY 465 ACACCTAATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 524
DB 72864 AAAAAATATATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 72923
QY 525 TCACATCTGACATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 584

```

[illegible]

| | | | |
|----|-------|-------------------------------------------------------------------|-------|
| QY | 1693 | AATATAAATAATAAAGCGTCAAAAAATATGAATATACAGCTCACCGCATATATAAAGACG | 1702 |
| Db | 74062 | TAATTTTAA-----TTTAAATTAATTTAAATATAATATATATATATATATATATATAT | 74111 |
| QY | 1703 | TCCAAATACCCATCGTTAACTCCACACTAGTAGTAAACATTTTACATACAACTTAAGT | 1762 |
| Db | 74115 | TATATTTTAAATTTAAATTAATAATATATATAATATTTAAATTAATTTTGGTTAAATTTAA | 74177 |
| QY | 1763 | ATAACTTAAATTTCCAGGTAAATTAATATATAACAAATTCCTATCTCCTATCTCATTTAC | 1822 |
| Db | 74175 | TTACATATTAATATATTTAAATTTATTTAAATATTTATTTATTTATATATTTTTTTAAATAA | 74233 |
| QY | 1823 | CAATTTGTTACAGTTTAAATAAGTCAATTTAACTTTTACTTATTTATAAATAATAT | 1882 |
| Db | 74235 | AAAT--ATATATTTAAATTAATATGATATATAAATTAAGAAATTAATATTTAAAAATATAT | 74299 |
| QY | 1883 | GTTTGCACTTTTTCACATCATCATGTATATATATGAACATAATTTTACATAAAATTAAT | 1942 |
| Db | 74293 | GTAATTTAAATTTATTAATTAATTAATATATATATATTTATTAATTTTATTTATATATATG | 74355 |
| QY | 1943 | AATTAATAAGAACCATACATATAATTAATTAATACATAAAACACCTTTTACTACGTAAATTA | 2002 |
| Db | 74353 | TTTATTTATATTAATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTTAAAAATTA | 74411 |
| QY | 2003 | CATGATATGCTCGACATATATAAATTCATGATCTCTTGACATATTTCCAAACATAAAAATATG | 2062 |
| Db | 74413 | TTAATTAATTAATTAATTAATTAATTAATTTATTTTAAATTAATTAATATATATATTTAATTT | 74477 |
| QY | 2063 | ATAACGTTTTTTTATAGCC-----ATTAGTCTTAAATTAATGCTAATTT | 2109 |
| Db | 74473 | AATATATTTTATTTTAAATTAATTAATTAATTAATTAATTTAAATTTATTTATTTATTTAT | 74533 |
| QY | 2110 | ACCCAAAAAGGTTAAATATACATATCTCTGCAACAATATGAAGAATACATACACGCTCA | 2169 |
| Db | 74533 | AATATATTAATTTATTTATTTTATTTTATTTTAAATTTAAATTTATTTAAATTTAAATTTTAA | 74599 |
| QY | 2170 | AATATAATAAACAATTTCTAAATTAACAACGACGACCTTAAACATAGAACCAATATATTA | 2229 |
| Db | 74593 | TTTATTTAAATATATTTATTAATTAATTAATTAATTTATTTTATTTTATTTTATTTTAAATTTAT | 74655 |
| QY | 2220 | TCATACCTAAAGTAAGATATCCCTACAGCATGATCAATATATTAAGAAAAATCGTTATGT | 2289 |
| Db | 74653 | TAAATTAATTAATTTAAATTTTATTAATTAATATATTTATATATATTAATTAATTTTATTTTAT | 74711 |
| QY | 2290 | ATATTTTAA-----CAATTAACAATTTCCAACTGCAACAACAACGCAAGTCTATTA | 2345 |
| Db | 74713 | TTTATTTTATTTTAAATTAATTTATTTATTTTCATTTATTAATTAATTAATTAATTTTATTTAT | 74777 |
| QY | 2346 | TAAACAATTCAGATTTATATATATCTTACCAACAACATAGTTATCAATATCTATTAATACA | 2405 |
| Db | 74773 | TAAATTAAGAAATCAATTTTGTGTGATTAATTTTAAATAATATTTTATTTAAATATTAATTTAA | 74833 |
| QY | 2406 | TCATATCTCTTTTAAAGAATGACTTCAACAAT--CATACATAAATATATATCATGCAAA | 2463 |
| Db | 74833 | ATATTAACATTTATTTAAATTTATTTTAAATAATCCATTAATTAATATATTTATTTATTTT | 74899 |
| QY | 2464 | TGTTTATTCGAAAACATTAACAATTCGAAGATTTTCAATTTA | 2506 |
| Db | 74893 | AAAAAATAATTTTATTAATATATTAATAATTTAAATATTTATTTATTTA | 74935 |

| RESULT | 26 |
|------------|---------------------------------------------------------------------------------------|
| LOCUS | AC004157 |
| DEFINITION | AC004157 Plasmodium falciparum 169546 bp DNA linear HTG 12-AUG-2000 |
| ACCESSION | AC004157 |
| VERSION | AC004157.8 GI:9797712 |
| KEYWORDS | HTG; HTG_PHASE1. |
| SOURCE | Plasmodium falciparum (malaria parasite P. falciparum) |
| ORGANISM | Plasmodium falciparum Eukaryote; Alveolata; Apicomplexa; Haemosporida; Plasmodium. |

REFERENCE 1 (bases 1 to 169546)
 AUTHORS Hyman, R.W., Fung, E.L., Qin, F., Rowley, D., Mao, J., Tamaki, T.,
 Kurd, O.B., Conway, A.B. and Davis, R.W.
 TITLE Plasmodium falciparum 3D7 chromosome 12
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 169546)
 AUTHORS Hyman, R.W., Qin, F., Fung, E.L., Conway, A.B. and Davis, R.W.
 TITLE Direct Submission
 JOURNAL Submitted (19-FEB-1998) Stanford DNA Sequencing and Technology
 Center, Stanford University, 855 California Avenue, Palo Alto, CA
 94304, USA
 COMMENT On Aug 12, 2000 this sequence version replaced gi:8810447.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 2 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

FEATURES
 source
 1. 23466: contig of 23466 bp in length
 23467 23666: gap of unknown length
 23667 169546: contig of 145880 bp in length.
 Location/Qualifiers
 1. 169546
 /organism="Plasmodium falciparum"
 /mol_type="genomic DNA"
 /db_xref="taxon:5833"
 /chromosome="12"
 /clone="PFYAC293"
 /clone="3D7"

ORIGIN
 Query Match 3.7%; Score 173.8; DB 2; Length 169546;
 Best Local Similarity 46.7%; Pred. No. 1.1e-13;
 Matches 1029; Conservative 0; Mismatches 1122; Indels 52; Gaps 13;

QY 345 AATATATAGAGTCTTCTTTTGGACATACATTTAAATCAATCAAAACAAATATAGT 404
 DB 84328 AATACATTTAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 84387
 QY 405 ACTATCTAATTAATAAAAAATCTCAATAGCTTTAGCAAGATCAAAATATATAG 464
 DB 84388 TTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 84447
 QY 465 ACACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 524
 DB 84448 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 84507
 QY 525 TCACATCGACATGATTAACATAGTAACATTTCTTAACCTGATATATACATTA 584
 DB 84508 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 84567
 QY 585 ATTAATTAATTAATTTGACGTCTTCTTGATATTAATTAAGTAAAAATGAA----- 636
 DB 84568 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 84627
 QY 637 AATGATTAATTAAGGCGACATCTTTGGCACACATAGCAATGACCTGACGCA 696
 DB 84628 AATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 84687
 QY 697 TGGTATATCACTTAATTAAGGACGATCATGAATATTAACATATTAACACTTA 756
 DB 84688 AATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 84747
 QY 757 CCGAAAAATTTAAATATTTTCCACATATCATCTACACACATATATGCAACG 816
 DB 84748 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 84807
 QY 817 CAATATAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 876
 DB 84808 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 84866

QY 877 AATATTCGACCATATGTATACAAATATCAACACAAACAGCATATGTAAAAATACTCC 936
 DB 84867 TAAATTAATTAAT-ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 84925
 QY 937 AATATTAATTTGGTAAAGATTCGAATTTACATCTCATTAATTTCCCGTGAATAGC 996
 DB 84926 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 84985
 QY 997 AGAAGAAAT-----TTTCAATTAATTCGAATTTTATTAATCTCTGTCGCAATTTGC 1050
 DB 84986 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 85045
 QY 1051 TAAAGAAATTCGACAAATCTCTCTTATATATATATATATATATATATATATATATAT 1109
 DB 85046 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 85105
 QY 1110 TTGGTGAATCTTGATAGAAACATATAGTA-ATGATGTTCTAAAAATATATACATCA 1168
 DB 85106 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 85165
 QY 1169 GATATTTATAGGTACGATTTGAGCTGCTTAAGACATTAATTGCA-----TTGCATG 1224
 DB 85166 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 85225
 QY 1225 TCGAATATCGCTGGAATAAATTTAGGTATTAATGCTGTTGCTACTGTTTAACTTAAG 1284
 DB 85226 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 85285
 QY 1285 CATGAATGAATTAATAACACTATATATAGCAAAAAATCAATTCATATGATCTTCATACATT 1344
 DB 85286 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 85345
 QY 1345 AATGCTCATCATGCTTGAGAGTCATATATTAATTAATTAATTAATTAATTAATTAATTAAT 1404
 DB 85346 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 85405
 QY 1405 TATGCTTTTGGACTAGAAATAGAAAAGGTAGAAAATTAATTAATTAATTAATTAATTAAT 1464
 DB 85406 TAAATGCTTATTTTATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 85465
 QY 1465 CCTAAGCTTGTGGAAGAACCGTACTGTGTAACACGATATCTTTGCAAAAAAAGT 1524
 DB 85466 ATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 85525
 QY 1525 TAAATGA--GAATACCAATTTGCATATCTATTTTAACATTTATTAATGATATATATAT 1582
 DB 85526 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 85585
 QY 1583 CTAGATTAAGCCATTAACAAATATATCTATCTATCATGCTTTGATTAATTTCAACAC 1642
 DB 85586 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 85645
 QY 1643 AATGATTAATTAACGGTCAAAAAATATAGAAATCAACGTTCCACCAATATTAATTAATTAAT 1702
 DB 85646 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 85698
 QY 1703 TCAATTAACCATGCTTAACATCACTAGATATGTAACATTTTACATTAACAACTTAAGT 1762
 DB 85699 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 85758
 QY 1763 ATAACTTAATTAATTTCTAGCTAATTAATTAACAAATTCCTATCTCAATTAAC 1822
 DB 85759 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 85818
 QY 1823 CAATGCTTACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1882
 DB 85819 AATAT--ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 85876
 QY 1883 GTTTCACACTTTTCTACTACATCATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1942
 DB 85877 GATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 85936
 QY 1943 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2002

[illegible][illegible]

[illegible][illegible]

| | |
|------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| RESULT | 28 |
| LOCUS | AEO15939 301278 bp DNA linear BCT 06-FEB-2003 |
| DEFINITION | Clostridium tetani E88, section 4 of 10 of the complete genome. |
| ACCESSION | AEO15939 AEO15927 |
| VERSION | AEO15939.1 GI:28203008 |
| KEYWORDS | |
| SOURCE | |
| ORGANISM | Clostridium tetani E88 Clostridium tetani E88 Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Clostridium. |
| REFERENCE | 1 (bases 1 to 301278) Bruggemann,H., Bauner,S., Fricke,W.F., Wiewer,A., Liesegang,H., Decker,I., Herzberg,C., Martinez-Arias,R., Merkl,R., Henne,A. and Gottschalk,G. The genome sequence of Clostridium tetani, the causative agent of tetanus disease Proc Natl. Acad. Sci. U.S.A. 100 (3), 1316-1321 (2003) |
| TITLE | |
| JOURNAL | |
| MEDLINE | 22457253 |
| PUBMED | 12552129 |
| REFERENCE | 2 (bases 1 to 301278) Bruggemann,H., Bauner,S., Fricke,W.F., Wiewer,A., Liesegang,H., Decker,I., Herzberg,C., Martinez-Arias,R., Merkl,R., Henne,A. and Gottschalk,G. Direct Submission Submitted (31-OCT-2002) Goettingen Genomics Laboratory, Institute of Microbiology and Genetics, Georg-August University, Griesbachstr. 8, Goettingen 37077, Germany |
| AUTHORS | Location/Qualifiers |
| TITLE | 1..301278 |
| JOURNAL | /organism="Clostridium tetani E88" |
| FEATURES | /mol_type="genomic DNA" |
| source | /strain="Wasschuselts" |
| | /sub_strain="E88" |
| | /db_xref="taxon:212717" |
| | 320..1207 |
| gene | /locus_tag="CTC00854" |
| CDS | 320..1207 |
| | /locus_tag="CTC00854" |
| | /codon_start=1 |
| | /transl_table=11 |
| | /product="oxygen-independent coproporphyrinogen III |


```

oxidase, Fe-S oxidoreductase"
/protein_id="AA035452.1"
/db_xref="GI:28203009"
/translation="MEEDSMRYEGVAVPSPSEANLLIQTTIGCSHNKSCFSCMYK
KRRVVELKEIFEDLEOGLYYKINRIPLADGALCKTESLKILKIKETHPOCD
RVGIYGTADILRKSIEELKEIKELGIGITWGLSGSDEILKINKKIVLPTFAS
KYKIGITLSTVVISISLGKKEKHEAIEFKAVSLIDPVLGVLTLLINEDTMAE
QVKGERTILKPEVWIEYTKLMDENADVNCIFSNHASNVPVPLKGTLPKDEKJLRT
IDNLSGDFYYKEDERLL"
/locus_tag="CTC00855"
/locus_tag="1518"
/locus_tag="1518"
/locus_tag="CTC00855"
/codon_start=1
/translation="hypothetical protein"
/product="hypothetical protein"
/protein_id="AA035453.1"
/db_xref="GI:28203010"
/translation="MSRNLIIILGGFGMNSNVYKSKSLKYLSPFIHLYVFLG
LVLYLKEKSNINHIYIMINIMWMSFISNIIIIKKEGGENLYPKKNAIITTSV
LSVVALIILICELIYFVKFTN"
/locus_tag="CTC00856"
/locus_tag="1943"
/locus_tag="1943"
/locus_tag="CTC00856"
/codon_start=1
/translation="hypothetical protein"
/product="hypothetical protein"
/protein_id="AA035454.1"
/db_xref="GI:28203011"
/translation="MEGSMNKSVYAFILMSCIISLIRKQVYLAQVSDPEQSLH
OELDRLYGEGNPSDEMLVFLNIGFDEKELIDYOSDARFNSKILMSGLROAN
RDQVYKNNKNNLLFAPSNRPGOKIESINPANIASLDGAGAGIAGCTPT
ALTSPAVIKAVIKSVGVVGSALGVANMISLRSNHTGRGKSEFYRGRTNDLYLD
NIIQGTIRITRY"
complement(2840..4396)
/locus_tag="CTC00858"
complement(2840..4396)
/locus_tag="CTC00858"
/codon_start=1
/translation="chemotaxis sensory transducer protein"
/product="chemotaxis sensory transducer protein"
/protein_id="AA035455.1"
/db_xref="GI:28203012"
/translation="WYKPILEGDVVNLSTLSKFORNTINVKSPAFGVNLIMIC
LFLVYDNFELFKLEIYAVSIIELILFLHKKIKSNKINIDNKNIGEMVITA
YNCIIPNTIPISNECLAIPEYMTFCLINKKIITVSAIPAPISEVILLILPSLI
PIGENIIMTAKTVIILINIGIYLSYSTANLIEKSNSEDEIKRNDKRTLLFNO
IKRISITLVISKDLITASEISANLQETISACVINOSDSWEDTKRKEITJQDL
YSSNQTSDRVVATGNFSKELVSIENKESNELISIIYSIKEGTTFPDKFLKXK
SNEIDQIAVINAVSEQNLINLASIEARAGEGKGFVAVANEIRLSNTKSLQ
NIAKINDKSLIDVENIMLNNTNITGVAILTDIYKELKTNINNNANVTNVEI
TEINKLKEIDTVVNLNEKISNGTELVDEKVTVAIENNTAVSEELNSABELND
IATNNSNLI"
complement(4390..5400)
/locus_tag="CTC00859"
complement(4390..5400)
/locus_tag="CTC00859"
/EC_number="2.3.1.41"
/codon_start=1
/translation="3-oxoacyl-[acyl-carrier-protein] synthase III"
/product="3-oxoacyl-[acyl-carrier-protein] synthase III"
/protein_id="AA035456.1"
/db_xref="GI:28203013"
/translation="WTKLRNVYIKSVGTVPVPHKKVHNKELIQHEQYGSQNEAKELM
DEIGETISMDENETVITMSLASKICLERAGLTPNDIVIASNTPELSPSTM
LIRDLNSKAGVLDVNCDCIOMGLQVATSLKLNQYKALVAGFGLSRGM
NISAATYADNSTAMILEVQEBEIRFGLGHSRFDSDQYIAIOFPACGMKIRSDQ
DEFEKRIADNNSHIFIPQSSADVTETLLTYNKPSPDYTOFPQSCSIEVNEITLNA
MSLPHNRFTPLADKGYTGPCSPPLALADRIQEKPFNESELCIMCSVAAGCITALLY
KM"
5833..6921
/locus_tag="CTC00860"

```

```

CDS
5833..6921
/locus_tag="CTC00860"
/codon_start=1
/translation="galactose-binding periplasmic protein
precursor"
/protein_id="AA035457.1"
/db_xref="GI:28203014"
/translation="MIMLISCKSGGVLMKKIISLYTAVALTTLFAGCGSKGKGGK
EGKELKGVLYKFDFTISTVDAIKENSIVKSKETGEKIVINAVDGGQOATQNDQ
VDTEITQGVIVAVNWVDRTAASVILKEKAKANI PVFENRBEVVDKMNKMYVVG
AKAEEGAMOGCIIADYKKNKPTADKNGKQYVWLDEGPGSDQIIITRESVKELE
KEGKIKTEKATDTANQKAGQEKSNMISAYDRIKAYFANNDDALNATIEKXAG
YKGDRTIPVGVDAITPAIDSLKAGBMGTVALDKAKOGKALFDIAYSIAKSEDPK
SVGEITDGKRYKVPYSITKENLKDYER"
7005..8519
/locus_tag="CTC00861"
/locus_tag="7005"
/locus_tag="7005"
/locus_tag="CTC00861"
/locus_tag="mgIA"
/locus_tag="CTC00861"
/codon_start=1
/translation="galactoside transport ATP-binding protein mgIA"
/product="galactoside transport ATP-binding protein mgIA"
/protein_id="AA035458.1"
/db_xref="GI:28203015"
/translation="WVENNIILEMNGISKPFVYKALDGVDLKYKKGTYHAALGNGA
GKSTLWKLCLFGIYRSDGELVIGKRVQFKNADALENGISMTIHSLHPVPRSVEN
VMGRPPVKKVFGIGVDRKQVEDTDLGKLNKINIDNTIVSKLSVSGVQGLIAK
AVYNSKIIYMDPTSLTENEVTHLFINISDLKQGVAI IYISHQKEELIKIADSVT
IMDGKII GTWEBAEGTTDLISKVGADLTNSPREKNTPGVIMKVENLITKADNS
FKDISERKELGIGTIGVAGARTLVSISFGLAKIEGKIYINQGVYKISPIKSK
KXIALTERBRSSTGIPUTLVGDNITIGDKYIDLKPYVNOKRGKMKIKSIECLN
IRTPSHATQKNSNGENQKVFIRSWLITBEDVLIWDEPTREGIDVAKKIYISIDSL
SKWKSIIIMTSSMPELIGSDRIMTICDGRSLGIIIEGEPATQSEIMKATATPI"
8539..9549
/locus_tag="CTC00862"
/locus_tag="8539"
/locus_tag="8539"
/locus_tag="CTC00862"
/locus_tag="mgIC"
/locus_tag="CTC00862"
/codon_start=1
/translation="galactoside transport system permease protein
mgIC"
/protein_id="AA035459.1"
/db_xref="GI:28203016"
/translation="MEFVYKKEKINTKIKFAGKYAYVULALLLSIGFDSNPL
SKNIIINVLIAVRITLAVGAGALITGTDLISAGRVGFACIASILLQRPDVAIK
LFPNAQLPILIPILVITIGLIVINQVAVYLVPPFATLGMVIVYGAACVY
NAQPIGLKDLFESLGTSLGIPNIIIGAVIIVWVILINKTKFGKHIYAGCNP
AAVSGVNDVLTIVKYALAGALGALLAARTGANNGLIMELDAIAAATIGG
VSTSGGIGVQGIITVLFVFNNGVLVLSAVAWQVYIKGIIIGAVAFDIRKXIA
KX"
9646..10650
/locus_tag="CTC00863"
/locus_tag="9646"
/locus_tag="9646"
/locus_tag="CTC00863"
/locus_tag="laci"
/locus_tag="CTC00863"
/codon_start=1
/translation="transcriptional regulator of the lacI family"
/product="transcriptional regulator of the lacI family"
/protein_id="AA035460.1"
/db_xref="GI:28203017"
/translation="MMPIINDVAREGVSIISVSRVNNNYPVKETEIRKIEALEKL
NKPKNLALSIITKSTIGVLPGLTLPPTIVESIDYDQGVSYVLENTRBEA
LKEKEITELMKSQGDGIIIDPTLNLKXLYEDLSMKTPIIIVGSGSEELKNGIC
YBKCVGTLEAPVYLAIEGKEIATFRGSKSYVDIKEXYTDIIRKNTKXNEIDVG
KNSIIEVENTQKVEEFLNKKPAPAIACDMLAGVYVANCNKLNANVPDISIG
FNTLIISNIHPLTIVDLMKRIGNIALKLEIIEIDKNSIKYTMETKLIVRESY"
10734..12287

```


CDS

/locus_tag="CTC00864"

10734..12287

/locus_tag="CTC00864"

/EC_number="2.7.7.10"

/codon_start=1

/transl_table=11

/product="galactose-1-phosphate uridylyltransferase"

/protein_id="AAO35461.1"

/db_xref="GI:28203018"

EDVESDINATDIDNMLKYAIDNDIEDGTDKDIEDTRIMVLTPEASEIVRNFY

NIYKKGAEATNFYNAKALNRYCRIRKNSWKTIDYGMETINLSKPKDS

KDIARAKPOISYRCPCLLENVGRILHDPARONRIVPELIEEKWLOYSFVY

YKXHSILFEYKHEVPMKISNKTFRHFLFIEKFPYHFGSNADPIYGGSLINHEHQ

Query Match 3.7%; Score 173.4; DB 1; Length 301278;

Best Local Similarity 51.7%; Pred. No. 9,8e-14;

Matches 477; Conservative 0; Mismatches 431; Indels 15; Gaps 3;

QY 620 ATTAGGTAAATAATGCAATGATTTAGGTGAGATCATTTGGCAGCAATAGCA 679

DB 296990 AATATGACATTAAGGTATTAATCTTTATGAGAGAGAGATTGTGATCGTTTGGT 297049

QY 680 ATTGCACTGTGACACATGTGTATCATGTTAATTATGGGACGTGATCATAAGAAATTT 739

DB 297050 ATTATGTTAGCCAAAAGGTTATTAACATTAATATGGAATAGAAAACCCCATGTGTA 297109

QY 740 AACATATTAACACTTACCGAATAAATTAAATTTTACCAATATCATCTACCAAGAC 799

DB 297110 GCAGATATTAAGCAAAAAAATAAATATTAATTTCTCCCAATGTGTTATTCATCA 297169

QY 800 AACATATATGACACAGCAATATGACGAAGTATTTATGACAAACATATGATTAATTC 859

DB 297170 AATGTAAAGCCTATTAAGGTATGAAGAAGCTTATGAGGGAAT---TAAATATGATGTT 297226

QY 860 TTAACATATCTCTACCAACAATTTAGCAACATATGACCAATAAGCAACAACAGAT 919

DB 297227 ATTCTGTACCATCTCATGCTTATAGAGAAATATTCAGAAATATG-----AAAGATTAC 297280

QY 920 ATGTGTAAATAATGCTCCAAATATTAATTTGTAGTAAAGGTATCGAAATTAATCATCTCAA 979

DB 297281 TTTAAAGAAAGATGCTATTTATATATAGTGTGCAAAAGGCTATAGAAACATATGAGAAA 297340

QY 980 TTCCAGGGAATATGCAAGAAATTTTACCAATTAATCAATTTTATCTCTCGT 1039

DB 297341 AGGTATATCTCAATATATTAAGAGAAATTAATCTTAAATCTCTGCTTATTCAGGA 297400

QY 1040 CCAAGTTTGTCTAAAGAAATTTGCAAGACATCTTCTGTAGTATAGTACTTGTGAT 1099

DB 297401 CCAAGTCTATGCAAGAAAGATAGCAACAATATTAACCACTGATGATGATTAATCAGAA 297460

QY 1100 AATAAGAACTGTGGATTCATTTGATTAAGAAACATAAATTAATGATTTCTTAAATAATA 1159

DB 297461 GATGTTAAAGCATCTTTTGAAGTACAGAAATTTATTTAGTACAAATAAATTTAGAGTATAC 297520

QY 1160 TACCATCAGATATTAAGTGTATACAGATTTGAGGCTGATTAAGAAACATATTTGCAAT 1219

DB 297521 ACTAATGATGATATATAGTGTGTAATAAGAGGGGAGTGAAGAAATATATTTGCTTTA 297580

QY 1220 GATGTGGAATTAATGCTGCAAAAAATTTAGTATATATGCTGTGTTGCTGCTTATTAAT 1279

DB 297581 GCTGTGTGATCTCGATGGAATAGAGATATGAGATTAATAAAGGCTCTTAATAGCT 297640

QY 1280 AAAGCATGATTAATTAACACTATATATAGCAAAAAATTAATCAATAGATCTTCAT 1339

DB 297641 AAGAGAAATTAATGATATATAA-----ATTGAGAAAAAGTTAGAGGACCAAGGGA 297694

QY 1340 AATATTAATTTGCTCATCTGCTTTGAGATCTAATATTAATGATGACAAAGAAATTTCA 1399

DB 297695 ACCTTTGGGGGCTAACAGGTATGGGGATATGATAGTAACTGATGATGATGATGAT 297754

QY 1400 CGCAATATGCTTTTGGACTAGAAATAGGAAAGGTAGAAATATTAATTAATTAATGAT 1459

DB 297755 AGAAATAGAGAGCTGTCTTTTATATGAAAAAGTTTATCTATGGAAGAGCTATAGAA 297814

QY 1460 CACAACTTAAGGCTTTGTAAGGAAACAGTACTGTATTAACCACTGATATATCAGCAAA 1519

DB 297815 GAAGTAGGATGGGTGTAGAGGATATTAAGAACTGTGTAAGCTTTTATGATTAAGGAA 297874

QY 1520 AACTTAATGTGTAATTAACCAAT 1542

DB 297875 AAGTTAAAGATATCTATGCCAT 297897

RESULT 29

AP003134/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 Kuroda, M., Ohta, T., Uchiyama, I., Baba, T., Yuzawa, H., Kobayashi, T., Cui, L., Oguchi, A., Aoki, K., Nagai, Y., Lan, J., Ito, T., Kanamori, M., Matsumaru, H., Maruyama, A., Muraikami, H., Hosoyama, A., Matsumaru, U., Y., Takahashi, N., K., Sawano, T., Inoue, R., Kato, C., Sekimizu, K., Hirakawa, H., Kohara, S., Goto, S., Yabuzaki, J., Kanehisa, M., Yamashita, A., Oshima, K., Furuya, K., Yoshino, C., Shiba, T., Hattori, M., Ogasawara, N., Hayashi, H. and Hiramatsu, K.

Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*

JOURNAL

Lancet 357 (9264), 1225-1240 (2001)

MEDLINE

21311952

PUBMED

11418146

REFERENCE

2 (bases 1 to 301550)

Director-General, Biotechnology Center, Aoki, K., Oguchi, A., Hosoyama, A., Nagai, Y., Kuroda, M., Hiramatsu, K. and Kanamori, H.

Direct Submission

Submitted (30-JAN-2001) Director-General, Biotechnology Center, National Institute of Technology and Evaluation, Biotechnology Center; 2chome 49-10 Nishihara, Shibuya-ku, Tokyo 151-0066, Japan (E-mail: bio@nite.go.jp, URL: http://www.bio.nite.go.jp/, Tel: 81-3-3481-1933, Fax: 81-3-3481-8424)

On Jun 12, 2001 this sequence version replaced gi:13701258.

COMMENTS

FEATURES

source

1.301550

/organism="Staphylococcus aureus subsp. aureus N315"

/mol_type="genomic DNA"

/strain="N315"

/sub_species="aureus N315"

/db_xref="taxon:158879"

/complement(249..947)

/gene="SA1293"

/complement(249..947)

/gene="SA1293"

/note="ORFID:SA1293"

/codon_start=1

/transl_table=11

/product="conserved hypothetical protein"

/protein_id="BA042553.1"

/db_xref="GI:13701259"

/translation="MSIETMIIIVFVILMTVPLMAOHKVKNSYKYSQVRSIGTKGR

EVALRIHANGIYDEYVKGEGFLTDHYDPKKVSLSPANDRSPVAGTAAAEVG

HAIDHOGYWFLEFPAALVPVANTLSSISYIMIGITLTAIGSAFSGTALMIGAGLM

SLAVFSTVTLFVERPDASSRAMKQITALNIVKEXKARVLSAAANTVYAAATVAV

AEVRIILIIARSDD"

complement(1001..1588)

/gene="SA1294"

/complement(1001..1588)

/gene="SA1294"

/note="ORFID:SA1294"

/codon_start=1

gene

CDS


```

12452 AATGCTAAATTTACCGCTTTTGGATGGAAGTTTGGGACACCCCTTGGCAATGTTCT 12393
688 GTACACACATGATATATAGTAACTTATGGGAGCGTATCATGAAATATATACATAT 747
12392 TCGAGAAATGACATGATGTTTGTATGCTGGGATTAATCAAGTGTGTATGATAT 12333
748 AAACATTAACGAAAAATTTAAATATTTAACCCACATATCATCTACCAACAAATATA 807
12332 AAATCATGCTATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 12273
808 TGCAACCGCAATATAGACAGATATATCTGACAAACATATCATGATATATATATATAT 867
12272 CCGTCTTCAATATGACACAAAGCATTCATTTGCAATATTTACTTAATG---CTTT 12216
868 TCCTACACAAATTAACGACCATATGATACAAATTAACAAACAAACGATATGTTAA 927
12215 ACCTAATAAGCATGCGGAGAGTTGCTACTCAATTAATGATAGC-----TGACCTC 12162
928 AAATCTCAATATTTTGTATGATGATATGATATGATATGATATGATATGATATGAT 987
12161 TAAAAAGCTTTTATACATGTTGCTTAAAGGATGAAATGGAAGCTTAAACGCTGTC 12102
988 TGAATAGCAGAGAAATTTT-----ACATATATATCAATTTTATATCTCTGCG 1038
12101 AGAATGATGATGATATCTATTTTACCGGATATATATGACAGTATGCGCGTGTGACG 12042
1039 TCCAACTTTGCTTAAAGAAATGCAAACTCTCTGATGATGATGATGATGATGATG 1098
12041 GCGAAGTATGCGGAGAGAGTTGATGACAGCAACCACTACAGTGTGATGATGATG 11982
1099 TAAATAGAACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1158
11981 AGATTAAGTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 11922
1159 ATACATCAAGATATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 1218
11921 CACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 11862
1219 TGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1278
11861 AGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 11802
1279 TAAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1338
11801 TCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 11748
1339 TACATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1398
11747 GACATTTTCTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 11688
1399 AGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1458
11687 TCGAATTTTACATTAAGATTAATTAATGATGATGATGATGATGATGATGATGATG 11628
1459 TCAACAACCTTAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1518
11627 TAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 11568
1519 AAAATCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1578
11567 AGAAAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 11508
1579 ATCACTAGTAAAGCATATCAACATATATGATGATGATGATGATGATGATGATGATG 1638
11507 CTCAGTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 11448
1639 ACACATGATTAATTAAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1666
11447 GATATTTTACATTAAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 11420

```

RESULT 30

```

AE015938/c
LOCUS      AE015938      303414 bp      DNA      linear      BCT 06-FEB-2003
DEFINITION Clostridium tetani E88, section 3 of 10 of the complete genome.
ACCESSION  AE015938
VERSION    AE015938.1
KEYWORDS  GI:28202752
SOURCE     Clostridium tetani E88
ORGANISM  Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
            Clostridium.
REFERENCE  1 (bases 1 to 303414)
            Bruggemann, H., Baumer, S., Fricke, W.F., Wierer, A., Liesegang, H.,
            Decker, I., Herzberg, C., Martinez-Arias, R., Merkl, R., Henne, A. and
            Gotschalk, G.
            The genome sequence of Clostridium tetani, the causative agent of
            tetanus disease
            Proc. Natl. Acad. Sci. U.S.A. 100 (3), 1316-1321 (2003)
JOURNAL    MEDLINE
PUBMED     24457253
PUBMED     12552129
REFERENCE  2 (bases 1 to 303414)
            Bruggemann, H., Baumer, S., Fricke, W.F., Wierer, A., Liesegang, H.,
            Decker, I., Herzberg, C., Martinez-Arias, R., Merkl, R., Henne, A. and
            Gotschalk, G.
            Direct Submision
            Submitted (31-OCT-2002) Goettingen Genomics Laboratory, Institute
            of Microbiology and Genetics, Georg-August University,
            Griesbachtstr. 8, Goettingen 37077, Germany
            Location/Qualifiers
                1..303414 "Clostridium tetani E88"
                    /organism="Clostridium tetani E88"
                    /mol_type="genomic DNA"
                    /strain="Massachusetts"
                    /sub_strain="E88"
                    /db_xref="taxon:212717"
                    88..1674
                    /locus_tag="CTC00580"
                    88..1674
                    /locus_tag="CTC00580"
                    /codon_start=1
                    /product="ABC transporter ATP-binding protein"
                    /protein_id="AAO35197.1"
                    /db_xref="GI:28202753"
                    /translation="MTSTNLSLRGGRVLEFDNVKFTPGNCYGLIANGAKSTPL
                    KILSGEIEANKGDVILEPGRKAVLKQHPQDSFVLETVIGERLNTNKEKEKI
                    YKEDPSEEDGKVALEBGEFTELNGWESDAATLLGLGIVELHYKMKDLGGQ
                    KYRVLLAQPLFGKPDILLIDEPNLDLSKIDWLEFLINENTVILISHRHFANKY
                    CTHLDLDFGKIQLVYGVNDFWYESSQALAEWAKQKKKEKIDLEFIFRFSANA
                    SKSKATSKKMLDKITLEDIKPSRRRPYVGFKEDREGNDLLAVEGKTIDEEKY
                    LNNVSRVYKGRKIAFVGENSRFARLQIILGENDKADSGEKGITTSQSYFEDMS
                    NTFNGVNLVLVLMKQFSESEKSETRIRGFLGMLSGEBALSKAIVISGGEVRCMLS
                    KMLSGANVLILDEPTNHLDESTRVANNGLINFGTILLFSDHQFQVITANRIEI
                    TPDGIIDRQMTYDELYNSKR"
                    complement (1752..2954)
                    /locus_tag="CTC00581"
                    complement (1752..2954)
                    /locus_tag="CTC00581"
                    /transl_table=11
                    /product="predicted amidohydrolyase"
                    /protein_id="AAO35198.1"
                    /db_xref="GI:28202754"
                    /translation="MKERRARLMLIIKQCNLNMADIYEERKDIENGKLIENIEI
                    DISKHDYKIIDAKNFVTPGVIPGVEPHCOLGIEEIHHPDGDSNEIDPILPOLAID
                    AINPDDEGDMARKAGVTTVTGTPRESNLIIGTTPAATGVTGNDMTVKDIAKPS
                    LGNTVNIITGKGNMFKTRWASAMITLIRLIGEFELNATVDCSAYLLPRLKYST
                    KLVIGAYGGRDPVENVNDIIGIMENNDNFALISSGHPENVNLELTVQITWYRK
                    GSRKVALKAMTINAKIILGLEDRIGSIEIGDADIILWNEPDLVYSYPTVLIDGN
                    VYVNEVI"
                    complement (2908..4116)
                    /locus_tag="CTC00582"
                    complement (2908..4116)

```

```
/locus_tag="CTC00582"
/codon_start=1
/transl_table=11
/product="predicted amidohydrolyase"
/protein_id="AA035199.1"
/db_xref="GI:28202755"
/translacion="MLIKNCLIDMVGIIHKEKYLILNKKRIKIEKEIDKOLENV
OVDAKGRITVPLGLIDPHCHMGVGTATLEGDGNETSPSTPELRAIDALDPWDVAY
DVAIGHVTVTVGPGSANVIGGFTAKTNGNDIKDCVKEVCOMKALGEMKVPY
GKQKAPKTRMNNALINRENLPKAEVSKKKEKKNESDONTSEFKYDIGNESLGR
VFDGRVKNHAKQSDIDITAIRICEPELRTITHCRTGYIIVDELKSKVQCITGPT
AGGSKLEIQHNEKLCARTFEBSGIDFALMTDNPVITIEGQWQALAPKNGSKETPA
LKGITINARLITDIDRVGSEIEIGKDADIVIMWDPIDTSEAGIVIIDQVYVERA
GEIDVYKLL"
complement(4188..5753)
/gene="mlrP"
/locus_tag="CTC00584"
/complement(4188..5753)
/gene="mlrP"
/locus_tag="CTC00584"
/codon_start=1
/locus_tag="CTC00584"
/product="putative efflux pump component mlrP"
/transl_table=11
/protein_id="AA035200.1"
/db_xref="GI:28202756"
/translacion="MAKSKKSLFDFRIFQVVEVAGNKLPTPTLFLVLIILICVLSV
LSMGVSVYVMAAGKAGAKMETTAVAVNLSSKQPSFLDFTKTEPEPLDLMV
MMGIGVEETGFINALMRKTLISAPSLITAVANLISADAMITPATIGMI
LFSKIGRNPIYIGITIGPAAGSGGFTANLFGVGTDLISGISESAAEARITPHPI
NYPWAKTAFVITVGVVTLIAEKYIIPKIDGSGSTEDCDLAAHLELDEKGLSNAGI
ATITITITLAIKSTPKNGLRADDTITPKPSLISGIVPLFFPLTQVYEGGITY
ITSEKDIPEKIMKNSLSGALPLVAVLPSITIKFPMESKLTVMKVSSELISKANG
VPLPLSFLVAVANLMTSGSSKMLLAPIFPMFAMNLSPASVQAVARVDSAT
NITSISSSLAVIGIMEQYNEKDEPVIGIVALTLPFALGILLSXSLILIMVTE
DPLGPNIFLS"
6191..6700
/locus_tag="CTC00585"
/locus_tag="CTC00585"
/codon_start=1
/codon_start=1
/product="hypothetical protein"
/protein_id="AA035201.1"
/db_xref="GI:28202757"
/translacion="MKKGYIAKRLKREEDIRORLYEBSLKKEISNVKHNPIIDPEA
NKSXVPTAETTFNNDCKKVESNTYVAELDEDGTAELPIATVAVYELDEBG
TLEIKKPIPYDICYCHLSDIRDSKYGEGIRLPWGINQFVIGLINGKILMSNFKDI
VEDISNKEK"
6801..7436
/locus_tag="CTC00586"
/locus_tag="CTC00586"
/codon_start=1
/codon_start=1
/product="hemolysin II"
/protein_id="AA035202.1"
/db_xref="GI:28202758"
/translacion="MGTNMPYKKEKEINAIHIGIVLPSIYAVLVAIVPAKTKDAM
YVVISIGSTILILYMCSTLYHSTFNEKVKIKFKPHSSIFLIIACTYPTPTITIL
RQMLGWSIIGIILWITVIGIVKIVCFEKEVKSFTFLIAGMWIVVAKLSISLSPV
KGIIVLIAGLIIYVGCIFPYAKDKIPLYNAIWHFVLGSGVCHPFSILLY"
7393..8022
/locus_tag="CTC00587"
/locus_tag="CTC00587"
/locus_tag="CTC00587"
/codon_start=1
/codon_start=1
/product="conserved protein"
/protein_id="AA035203.1"
/db_xref="GI:28202759"
/translacion="MEVYISPPYFTYFNYIKFISGCLILMLNLELIADIIRKGA
YAKDPLEHEPNKREYCKENLGHGPNMCPBGSEIEDLKKEVMPKRGIIYO
TIRPVKNTIRDRATDVIRDKHNTFTLELLENREKIIDLEFPKMGAGCELCCTTCGYL
```

```
KEEKIYPOKAISAEVGINLSLKSCLQFNYSDDSLAYVGLIK"
complement(8089..9645)
/locus_tag="CTC00588"
/complement(8089..9645)
/locus_tag="CTC00588"
/codon_start=1
/transl_table=11
/product="ABC transporter ATP-binding protein"
/protein_id="AA035204.1"
/db_xref="GI:28202760"
/translacion="MSIISVNTLTHGFDRALFNDVSENLKGEHIGLVANGEGKSS
PMIITGNLPPBPGVIVEMSKRYVGLDHTYLLKSGTTRERILKTAFLPLLEKEM
EIDCKYKANSDBLELLELDGFIQDILTNDFYIIDAKIDVYAGKAGADIGLRNDV
QDLSGGQRTKILAKILLERPDILLDEPTNLDGQTEMLKRYIQDENATLISHD
IFPLNSVINLIYHMEGELNRYVGVYDPLKYKAKQQLAEAYKQOQELQELKDFV
SRKARVSTRNAMSROKLDIDMDLIELPKDKPEFNEKRGTSKMLFETKDLVG
VNEALSPLNKHMERGKTAIVANGALIKSLIGIKRISGALTGLDINIGYF
EKEIKYANNCTIEYWSFSPFTQVEVAAALAKGLTTRKHEKVLVLSGEOAKVR
LCLINKETNPLVDEPTNHLVDVDAEELKRAIKAYKSGVLLICHEPEFYQDIIYDVW
NCENWTKIV"
9844..10482
/locus_tag="CTC00589"
/locus_tag="CTC00589"
/codon_start=1
/transl_table=11
/product="conserved protein"
/protein_id="AA035205.1"
/db_xref="GI:28202761"
/translacion="MDKNIEWYIDKVCRTIENLEKNIIEAFVYDEDELKKNIDEMI
PKGIVTAGDSMTLETVGIDMLRNDKYIYDKHKEGITESEKNTYNSFTAHYLS
STNALTEGSELVINDGSRVAMITYGPOVIVAGINKIVNLEAEKRVNVAFL
DARLRKNTPKAKIGYCVDCSSERICNFTYIKQFKRITVITVNNKLG"
complement(10604..11065)
/locus_tag="CTC00590"
/complement(10604..11065)
/locus_tag="CTC00590"
/BC_number="1.15.1.1"
/codon_start=1
/transl_table=11
/product="superoxide dismutase (Cu-Zn)"
/protein_id="AA035206.1"
/db_xref="GI:28202762"
/translacion="MAFAYCAIKGGLAPISIRGMYFTDPGGCKYCTITINGLPPYKP
ASGNMPVPGHGFHIEFGNONGDPEKFSKGMMPTNCPHAGDAPFVIFSN
GYSFMCPTNKPEVKDILKSLIHSPPDYNTQPGDSGRRLACGIVRV"
complement(11185..12345)
/locus_tag="CTC00591"
/complement(11185..12345)
/locus_tag="CTC00591"
/codon_start=1
/transl_table=11
/product="conserved protein"
/protein_id="AA035207.1"
/db_xref="GI:28202763"
Query Match
Best Local Similarity 3.6%; Score 172.8; DB 1; Length 303414;
Matches 517; Conservative 0; Mismatches 532; Indels 5; Gaps 2;
QY 2395 TCATTAATACATATCTCTTTTAAAGATGACTTGCACAAATTCATACATAATTA 2454
Db 127005 TTTCGAATGACAGATMAAAAGTGTGTTTCCTTACATTTGACTTCTAATGTGTT 126946
QY 2455 TACATGCATGTTTATTCAGAAACATTCACATCAAGATTTTCATTTAGCGGTAA 2514
Db 126945 TTTTGTACTTGTTCCTTACATCTTTACATATATGAAATTTTCAAAATTCATATTA 126886
QY 2515 TGCTGAAGTTCATTGCAACAAATTCAGAGTTCGATACATTTACCTGCAAAAGA 2572
Db 126885 TGTTTACCTATCTTTTCATCAAGAGTTCATATATATGAGCTTTCAGTCTGAAT 126826
QY 2573 GATAACGAGTGTTTTAAATACGTCTTCAAAATCTTCACAGACCTTACGTAATGTGTA 2632
```

| | |
|-----------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| TITLE | |
| JOURNAL | Lancet 357 (9264), 1225-1240 (2001) |
| MEDLINE | 21311952 |
| PUBMED | 11418146 |
| REFERENCE | 2 (bases 1 to 346900) |
| AUTHORS | Ohta,T. |
| TITLE | Direct Submission |
| JOURNAL | Submitted (28-FEB-2001) Toshiko Ohta, University of Tsukuba College of Medical Technology and Nursing, Department of Medical Technology; 1-1-1 Ten-nodai, Tsukuba, Ibaraki 305-8577, Japan (E-mail:tohaka@tsukuba.cc.tsukuba.ac.jp, Tel:81-298-53-3454, Fax:81-298-53-3454) On May 29, 2001 this sequence version replaced gi:13875626. |
| COMMENT | |
| FEATURES | Location/Qualifiers |
| SOURCE | 1..346900 |
| gene | /organism="Staphylococcus aureus subsp. aureus Mu50" |
| CDS | /mol_type="genomic DNA" /strain="Mu50" /sub_species="aureus Mu50" 12..263 /db_xref="taxon:158878" /gene="SAV1313" 12..263 /codon_start=1 /transl_table=1 /product="hypothetical protein" /protein_id="BAB57475.1" /db_xref="GI:14247084" /translation="MLTITKIKEDQVIIVYIPEDDISNGKGSVTFFNNKDAAVIDES LSLEIDENBYFMLTRNKSFSVVRDFIEKQEPFNYKIATW" 987..1181. /gene="SAV1314" 987..1181 /codon_start=1 /gene="SAV1314" /codon_start=1 /transl_table=1 /product="hypothetical protein" /protein_id="BAB57476.1" /db_xref="GI:14247085" /translation="NMNQSYIGLVKLTLTNKKILIGKVLDIPDKVDNPDGNSIID TGRIYDISENKIKITYLQK" complement(1725..2750) /gene="SAV1315" complement(1725..2750) /gene="SAV1315" /codon_start=1 /transl_table=1 /product="conserved hypothetical protein" /protein_id="BAB57477.1" /db_xref="GI:14247086" /translation="MISFENDYLEGAHEKYLLKLVDTNLVQASGYGDOPTAQAIEXKI KTDIDCNARIRPLVGOTNOGVINSMSLEYGSADTGHVAVEHGGAIBYGSHVCY ITIPSEKGASASDVETWETPFKSDPKDMVPGMWYISHPIREGLLYLSELEELC KKVKQOLP.FMDGARLGVMDSODMTFLAKCDVYDIAGTKIALCGEAIVFTPL KNNEPQPTFRIRGHGALLAKGRLLTCIQEFLETDPNIYISHAIMANMKDGFGP KGRRLFDSPFNQQPFIILSNKEIAELEQKVPKPAWEKTYDQHKVRAFAISMATTEML NKLLELI" |
| gene | 3020..3217 |
| CDS | /gene="SAV1316" 3020..3217 /codon_start=1 /transl_table=1 /product="hypothetical protein" /protein_id="BAB57478.1" |

```

/db_xref="GI:14247087"
/translation="MEQIKLKTFTAEELLEKNINFLSSEEAATNKLVTITKEIE
ERTPNNEEFPAITLSVNK"
3238..4755
/gene="SAV1317"
3238..4755
/gene="SAV1317"
/codon_start=1
/translation="cardiolipin synthetase homolog"
/product_id="BAB57479.1"
/db_xref="GI:14247088"
/translation="MFVFPFPGKTPMTPSNDLGTFTLLAGFTINLVAFII
FLERNRTASTWMLFVLPPLIGFLLPFGRTVSAKLNKNGNVLDPGLK
QOIESFDKNGTGNQVQKHDLVSKMLMDQDELNNKVDH IDGNLDYDLD
KNAXEYHLEYYTALDGLKRIHLAEKLEKQLEKILYDVGSKVMAFDF
KSLGEVAFASKLPLNFRNNNNHRIIVIDQGVGPNIGDEYGLGYW
RDLRIODGAVDALQRLFLDMNSQAHKPOEYDVKKFKNGPLGNSAAGP
ASDMHIEGYTKMTMSAKSVLOSPEYIPDNGYINAKIAKSGVDVHLMCKED
HPLVYATFNSASDLSSGVKITYTENGFHSKMLIDDEIVSVGTAMPREELN
ENAFVDEMLAKDLRAVEHDITKSQOLITESTANRPLSVKPKSLKVSPTL"
4939..5436
/gene="SAV1318"
4939..5436
/gene="SAV1318"
/translation="ATP-binding protein"
/codon_start=1
/translation="ATP-binding protein"
/product="ABC transporter homolog"
/product_id="BAB57480.1"
/db_xref="GI:14247089"
/translation="WIOSINIKSPNKKCVLKNISPDIEGKCIALGKNGAGSTLI
DILGNVANGSEIFDKDLQSENRSTIMPKMTPOPKVIEIINLYGSEYENPLT
EIIILTFDSSQLNQFNKLSGGQRLDVLISLIGQPOLILDEPTSTMDITREY
FMSII"
5461..5838
/gene="SAV1319"
5461..5838
/gene="SAV1319"
/translation="ATP-binding protein"
/codon_start=1
/translation="ATP-binding protein"
/product="ABC transporter homolog"
/product_id="BAB57481.1"
/db_xref="GI:14247090"
/translation="WLTSHYIEEYERWSDKIILLENGEIILNDSHIRTNOOST
LSDEYIKLKDLDVYQKHNNGTITKITSNVDTLIYLOQLHINDDIEIQNSIV
DSIFNKKQSGNSYDKLENNKI"
5807..6538
/gene="SAV1320"
5807..6538
/gene="SAV1320"
/translation="similar to ABC transporter integral membrane
protein"
/codon_start=1
/translation="similar to ABC transporter integral membrane
protein"
/product="hypothetical protein"
/product_id="BAB57482.1"
/db_xref="GI:14247091"
/translation="WLSYKIEBFVIRKKTLLISLFPVIFYILFTSLIEPEV
KPKFYKEMYSVTYSLSLSPILTFPDLINENKONEMQRMATPPEFTSYISKYK
TMLQRAILVIVGHTYKGVASAVOMESGFLWIGASLLITGILBSLMDIK
TSALANIVITGLAVIGLMPINTFPNVLQVAHVLPYHLRKLGVDIASNHNILIS
FAILLVYLGSIAYVCISHFKRAE"
5542..7633
/gene="SAV1321"
5542..7633
/gene="SAV1321"
/translation="similar to two-component sensor histidine kinase"
/codon_start=1
/translation="similar to two-component sensor histidine kinase"
/product="hypothetical protein"
/product_id="BAB57483.1"

```

```

/db_xref="GI:14247092"
/translation="MKFLKDSIAEISSILYIFPIAGIFENEVYGPKNLYISLVF
SLSYLLVIVNRLNTLMFYILLIHYFICFVSVHMLSLFPEYSAFVPEKN
NYKATNLFILMTICTIITYLNNVAVAVVYVSLIMLDFEKKRQREYKE
IAEKRNHTLLIAOEPRHIGODLHDITGASISIKSEILAYKILDTVEVKAELL
AINKLSRESLNKREIIDVYKPSFIEIDISIRKILKODIDPFTENKRLAOVSPK
QSLVMTITREAINVTKHAKSKVHKLKTVNNHKLMLTBDGKSIDCEVKSISQ
RVQHLNGTLAVDSITGFKIIEISTGILA"
7630..8232
/gene="SAV1322"
7630..8232
/gene="SAV1322"
/translation="similar to two-component response regulator"
/codon_start=1
/translation="similar to two-component response regulator"
/product="hypothetical protein"
/product_id="BAB57484.1"
/db_xref="GI:14247093"
/translation="WLSLIADQNMTRQAVQLIKHGDPEILADTONGDLAMKLE
ENRNVVITDIEMPGKTELEAIEIRKHLNIVITTFPRQYPERAYVVDYAV
LERSIEBELVETINKVNGEKSATLTSFVDKPLTFKQIVLREINGLSKREI
SEKLFITDGTVRNYSVLIIDKLPADNRFDMKRANEKWI"
8232..8533
/gene="SAV1323"
8232..8533
/gene="SAV1323"
/translation="ATP-binding protein"
/codon_start=1
/translation="ATP-binding protein"
/product="hypothetical protein"
/product_id="BAB57485.1"
/db_xref="GI:14247094"
/translation="WNISNENVTLLAVIAIIIGFIQIFPKLPLIVAVLSILGI
PVGFIYLVDSVFNKRKN"
8672..9205
/gene="nuc"
8672..9205
/gene="nuc"
/translation="thermophilic protein"
/product_id="BAB57486.1"
/db_xref="GI:14247095"

```

Query Match 3.6%; Score 172.8; DB 1; Length 346900;
Best Local Similarity 50.8%; Pred. No. 1.1e-13;
Matches 532; Conservative 0; Mismatches 492; Indels 24; Gaps 4;

| | | | |
|----|--------|--------------------------------------------------------------|--------|
| QY | 628 | AAAAATGAAATCATGATATTTAGCTGACAGATCATTTGGCAGACGATATGCAATTCGACT | 687 |
| DB | 199484 | AATGACTAAATATACCGTITTTGGTATGGAAAGTTTGGGACAGCCCTTGCAATGTCT | 199425 |
| QY | 688 | GTCAGACATGATATCATGTTAATCTTATGGGACGATGATAGAAATATTACATAT | 747 |
| DB | 199424 | TGCAGAAATGACATGATGTTTATGATGCGGATAAATCAAGATGCTGTTGATGAAT | 199365 |
| QY | 748 | AAACACTTACCGAAAAATTTAAATATTATCCCATATCATCTTACGACACATATA | 807 |
| DB | 199364 | AAATACATGATACATCAATTAATAAGATTTAAATACGGAATTTAGATTTACATCAT | 199305 |
| QY | 808 | TGCAACACGAAATATGAGAGATATTATCTGACACACATACATGATATCTTAATAT | 867 |
| DB | 199304 | CGTACTTCAATATACCAAGACATTTCAATTTGCAATATTTACTTAATG--CTTT | 199248 |
| QY | 868 | TCCTACACACATTTAGCAGCATATGATACACAAATATCAACAAACAGCATATGTAA | 927 |
| DB | 199247 | ACCTACTAAAGCATGCGAAGATTGCTACTCAATTAATGATTAAGC-----TGACCTC | 199194 |
| QY | 928 | AAATACCTCAATTTAATTTGATTAAGTATCGAATTTACATCATCTCAATTTCCCG | 987 |
| DB | 199193 | TAAAAAGACTTTATACATGTTGCTTAAAGTATGAAAAAGGACGTTTAAACGTGTCTC | 199134 |
| QY | 988 | TGAATATGACAGAAATTTT-----ACATATTAATCAATTTTATATCTCTCG | 1038 |

Db 15765 CTCATTAGCATTCAAAAAGCCCTATTTTATGTCATAGATCAGTAA-----AGC 15818
 QY 3199 ACTTTCCTCCTACCAAAATATTTAACCAAGTATTTAGTGGTATGATGAGTA 3258
 Db 15819 ACCCTTTTCAAAACCAAAAAGCTCATTCAATCAAAATTTGACGAAAGGCTGCATATT 15878
 QY 3259 CATGATATTAATGAGATTCATACCCCTGGATTTTATGATTTAGCTAGCAACTAC 3318
 Db 15879 AATGAAATTAAGGCTTATATGAGCGTTTGAATGAGTGAATATACAGAAATATAC 15938
 QY 3315 TTCTTTTCCACACCTGGCGAAACCAATATGATGATTCATCTGATGATGATGCTTT 3378
 Db 15939 CTCTTTTCTTCTCCTCCTTCTCAAAAACCTTACAGAGGATCAAGTTTTCAGCGCTTT 15998
 QY 3379 ATTAATCATCTCTTCAAAATTCAGTATTAACAGGAGATTCAGCTATTTATATCTC 3438
 Db 15999 TTGACTTAATTAATGATTTTCTTAATTAAGGAGATGAAAAAATCTTTATTTTC 16058
 QY 3439 AAATGCTGATTCACATCATCTATTTCTCTACGTAATCTACCAACTCTA-----TAGC 3492
 Db 16059 TTCTTTTCTTCTATTTTCTTTTCTTTTATCGGCTGTTTCTCTTATGATTTTACG 16118
 QY 3493 TCTCTTCAACTACTTATCTTCTCTCTCTGTAAGGCTTTCTATATATCATTAAGC 3552
 Db 16119 AGCGTTAATAGCTTGACTTAAGTATCAACATCAAAAGGCTTGTTAAATAATCTTTAC 16178
 QY 3553 ACCCATATGACAGACTTACAGCAGTGGCATATTTACATGCCCATATCATATATAC 3612
 Db 16179 TCCCAAGCAACCGCTTCTATAGCAAGGCTTAAGTACATTTCTCTCATATATATATA 16238
 QY 3613 AGGCAATATAGATACCTTTCTTAAAGCTTTCCAGTACACTTATCCATCATATATGCA 3672
 Db 16239 ATCATATTTATTTTCAAGGCTTTAAACAATTTATGCGCATCAATTTCTGCGATATTTAT 16298
 QY 3673 TCCCTTAACCATATATC 3690
 Db 16299 ATCTGTAATATCAATC 16316
 RESULT 33
 AR377574/c 1425 bp DNA linear PAT 18-DEC-2003
 LOCUS AR377574
 DEFINITION Sequence 2580 from patent US 6605709.
 ACCESSION AR377574
 VERSION AR377574.1 GI:40080756
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1425)
 AUTHORS Breton, G.L.
 TITLE Nucleic acid and amino acid sequences relating to Proteus mirabilis
 JOURNAL Patent: US 6605709-A 2580 12-AUG-2003;
 FEATURES
 source
 1..1425
 /organism="Unknown"
 /mol_type="genomic DNA"
 ORIGIN
 Query Match 3.6%; Score 170.6; DB 6; Length 1425;
 Best Local Similarity 50.1%; Pred. No. 1,5e-12;
 Matches 484; Conservative 0; Mismatches 474; Indels 9; Gaps 2;
 QY 2708 ATCATCTCCTTGAAGATTTCAATTAATTCATCTATTAACATTCAGTATGCTT 2767
 Db 1136 ATTAATCTCTTGGCTAGCGCTCATACCGTTAACACACATACATCTTTCTTAATGCTG 1077
 QY 2768 AAGTTACCTGGCATTCAATGATGATGCTGATTAAGCTTATCATCATTAATATCA 2827
 Db 1076 ACGTTACCTGACGAGTGGATTCATCATGCAATGTAACCTTGTGCTGAGACATTTG 1017

QY 2828 GTACAAAACCTATTTTAAAGATGCTATTCATTAATTAACCTACACAATTCGGTATA 2887
 Db 1016 GTTCAACCAAGCTCTGAGGCTTTTGTGAGAAATATACGGCTAGCTGGAAAT 957
 QY 2888 TCTGTAATATTTCTAATAAGGCTACTTAATTTGAAGACATTTATCTATATAT 2947
 Db 956 TCTTCTGTTCTATCAGCAGAGTGGGGCAATTAATACGTAATTAATGATGATGATG 897
 QY 2948 AATTCCTCAAAAACCTACAGCTTTTACTTCACTTCAATATCTTTGGAAGAAGACA 3007
 Db 896 AGATCTTCAAGAAATACACTTCCATTTAGCGTTTTTCTTAAGCTTTGATGATCCGCGA 837
 QY 3008 ATATTTCAACATCTATTAACAGGAATCTTAATTTTCCCTATATATTTTCCCTCC 3067
 Db 836 ATATATGACATCTGACTTTCACAGGGGATTAACGCCCAACGATTAATATGCTCTTCA 777
 QY 3068 TGTATATCTGATGATCTTAATGCGATGATACGTAATTCGTAATCTTCTCATCTATA 3127
 Db 776 GCCAAGACGGTATATAGCGGCTTGAATATCAAGTGCATATCCCAATTTCAATCAGCA 717
 QY 3128 AATTAAGTACCATGATTTGCTCTCTATATTTCCAAATATGAGAGTACTATAGAC 3187
 Db 716 AAGAGGAACCGCATTTGCTTCTTCAAAACGGCTTGCGTACTGATGATGACACAGTA 657
 QY 3188 AATATATATATCTTCTTCTCCTACACCAATATATTAACCAATATATTAAGCTGTAC 3247
 Db 656 AATGCCCTTTTCAAGACCAAAAAGTTCTGATTCATTAATGCTTTGGAAATAGCGCC 597
 QY 3248 ATAGATGATGATATATTAATGAGTATCATACCCCTGATTTTATGATATAGC 3307
 Db 596 ATA-----TTAGTGCATTAATAAAGGGCTTGTGCTCTGCGATATGGCGATGAAGACA 543
 QY 3308 CTAGCACTATCTCTTTCCAAACCTGGCGAAACCAATATGATGATGATGATGATG 3367
 Db 542 TGTGCAACTTAATTTCTTACAGATACCGCATGCGCATGATATAGACACATTAATGATG 483
 QY 3368 GTAGCTGCTTATTAATCAATCTTCAATTAACGATATGATGAGGAGTATCCGATAT 3427
 Db 482 CGAAGAAAGCGCCAAATATGCGATTAACCTCTGATGCGAGGGCTTGGCAATCAG 423
 QY 3428 TCATA--ATCCTCAATGCTGATTTCAATCATCATATTTCTCTACGTAATCTACGAC 3484
 Db 422 TCAGACACAGATGAAGGTAGTTCTGTGATTAAGTTGTTTGTCTCGAATAGTGC 363
 QY 3485 TCTATGCTCTCTTACACTAATTTATCTTCTCTGTAAGGCTTTCTATATA 3544
 Db 362 GTATATGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 303
 QY 3545 TCATTAAGCAACCATATGACAGACTTTACAGCAGTGGCAATATTAACATGCCACTATC 3604
 Db 302 TCAATGCGCTTGTGATTAAGATTTACCGCGCATCAAGATCAAGATGCGCTGCATG 243
 QY 3605 ATATTAACAGGCAAAATAGATACCTTCTTAAAGCTTTTCCAGTACATTAATCCATCA 3664
 Db 242 ATATATAACGGCAACGTAGGTTGCTCTTTAAGCTTTTAAATGAGATTAACCATCA 183
 QY 3665 ATATCAG 3671
 Db 182 ATATCAG 176
 RESULT 34
 AE007680 10861 bp DNA linear BCT 27-JUL-2001
 LOCUS AE007680
 DEFINITION Clostridium acetobutylicum ATCC824 section 168 of 356 of the
 complete genome.
 ACCESSION AE007680 AE001437
 VERSION AE007680.1 GI:15024666
 KEYWORDS
 SOURCE Clostridium acetobutylicum
 ORGANISM Clostridium acetobutylicum
 Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 Clostridium.

| | |
|-----------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| REFERENCE | 1 (bases 1 to 10861) |
| AUTHORS | Nolling,J., Britton,G., Omelchenko,M.V., Markarova,K.S., Zeng,Q., Gibson,R., Lee,H.M., Dubois,J., Qiu,D., Hiltl,U., Wolf,Y.I., Tatusov,R.L., Sabathe,F., Doucette-Scamm,L., Soucaille,P., Daly,M.J., Bennett,G.N., Koonin,E.V. and Smith,D.R. |
| TITLE | Genome sequence and comparative analysis of the solvent-producing bacterium Clostridium acetobutylicum |
| JOURNAL | J. Bacteriol. 183 (16), 4823-4838 (2001) |
| MEDLINE | 21359325 |
| PUBMED | 11466286 |
| REFERENCE | 2 (bases 1 to 10861) |
| AUTHORS | Childress,D., Zeng,Q. and Smith,D.R. |
| TITLE | Direct Submission |
| JOURNAL | Submitted (24-JUL-2001) GTC Sequencing Center Production, Finishing, and Bioinformatics teams, Genome Therapeutics Corp., 100 Beaver Street, Waltham, MA 02453-8443, USA |
| FEATURES | Location/Qualifiers |
| source | 1..10861 /organism="Clostridium acetobutylicum" /mol_type="genomic DNA" /strain="ATCC 824" /db_xref="ATCC:824" /db_xref="taxon:1468" 96..476 /gene="CACI702" 96..476 /gene="CACI702" /codon_start=1 /transl_table=11 /product="Hypothetical protein" /protein_id="AAK79668.1" /db_xref="GI:15024667" /translation="MNFTLMKKKKKKREKVSAIVRIAPVLTGLIIISMAADGIIYYEA VTIYIKNSNMKVVSEIYNPFKLVLQNEAKNLSDESYLRKIITYNNKNISTDEM SMLKNLKGDSKEVENIFPAGQL" |
| gene | 473..823 /gene="CACI703" 473..823 /gene="CACI703" /codon_start=1 /transl_table=11 /product="Methyl-accepting chemotaxis protein (fragment)" /protein_id="AAK79669.1" /db_xref="GI:15024668" /translation="WTSKPVIITPANPIDDGNTLVGVKTIIPDYFSKRPSDFKWG NSFKLSLNAAIEASRLBEGRGFVGVASEIKSLNNVETIIVIGIVAQINEKYVN MKDKTSTLNRDXYD" |
| CDS | 824..1039 /gene="CACI704" 824..1039 /gene="CACI704" /codon_start=1 /transl_table=11 /product="Hypochemical protein" /protein_id="AAK79670.1" /db_xref="GI:15024669" /translation="MAAYNKTKTNVNFSLQEIINNKKIVDENANVSSQNKNHYH EFRNISAFDEFNYSTLEENIGKFRLL" |
| gene | 1188..2084 /gene="CACI705" 1188..2084 /gene="CACI705" /codon_start=1 /transl_table=11 /product="Periplasmic phosphate-binding protein" |
| CDS | /protein_id="AAK79671.1" /db_xref="GI:15024670" /translation="MKSKKIKLMSVAIMTTIAGLPVCGSNSNSSSKSSSTTKKVS GSITTSALGPLAEKSVGEKEYPVSVVSGSGGTGNOLGAVENGSDIF AERKLNADDKALTIDHKVCAIGFAVTNNDVVKNLTKQVODIFGTKITNNVGGP SEAINVHPKXSSTRATPTKTVMGCKXKDGIQTGQASVSKTAMTSKAGSI SYLA FSLVTDEKDOIINISLDGVEATISKNVANGTYTPWASTEHHTYTGGKNDAAKAIDYM MSSDNALLERNGYIASNTRK" |
| gene | 2183..3073 /gene="CACI706" 2183..3073 /gene="CACI706" /codon_start=1 /transl_table=11 /product="phosphate permease" /protein_id="AAK79672.1" /db_xref="GI:15024671" /translation="WEKASFUKKUTRYVGRGFATICGVPIVILLTISITFPASKGS ITFKKHYSIFPLMSITWPKDPSTPKLGAFITLEGSTMVSGIAVILSPVALAFLV NIISSPKLDVRPKPSLELPGVIPSVVGMGLVTLVPFKRFGVSPGAAILAFV IMIIFTLASLSDAIKNVHNIEYGSGYLGTARMQIYKIIIPSSKNGILTGLGIA RAFGALAWOVIGIVTVSAQGLTFSPITLTSVLTMMANTPNQTAWDALWSLAFLL LVISITFIIVIRIFGRKEV" |
| gene | 3073..3957 /gene="CACI707" 3073..3957 /gene="CACI707" /codon_start=1 /transl_table=11 /product="permease component of ATP-dependent phosphate uptake system" |
| CDS | /protein_id="AAK79673.1" /db_xref="GI:15024672" /translation="MKSLEYDKIATIIILYLSLFVILLAFLIIFIYGRSDNLNLF LGNPRIAEKSGIGIGPELNFSPVLYLSLITVIGIAGIYSEYAKERINLFIIL CIEMASHRSIVGLFGLVFTVTHMGVITLSCALVITLINPSMTRENAIRASS SKVHASIGIGATNMOTIKNIVLPSAMELITGLIISAGRIFGBAALVLTAGMASFN INFDTISVDYDTSNPSLPRPAETLAHHIKWLNSGMTPDATKIANSSAVLVMTLFF NFMAITGRKITSYLSGK" |
| gene | 3959..4718 /gene="CACI708" 3959..4718 /gene="CACI708" /codon_start=1 /transl_table=11 /product="ATPase component of ABC-type phosphate transport system" |
| CDS | /protein_id="AAK79674.1" /db_xref="GI:15024673" /translation="MGIIOTKDLNLYGNVOALKINLDEPSANTVTLIIPSGCGKST FLRTNRMNDLIISVTKIDGEMFERGKVYDYDIELRKRVGMFOKPNFPMISYDN VAYGRIRIGIKRKKLBIDYERSKSAIEYVDRKSKALGISGCCOGLIARTL AVEPFLVMDPEPISALDPISTLKTEBELMDELKAKTYITVITHNQAGRISDNTAPFL NGEVENVKGTEDIYPKPKREDYITGRFG" |
| gene | 4750..5409 /gene="CACI709" 4750..5409 /note="phoU" /gene="CACI709" /codon_start=1 /transl_table=11 /product="phosphate uptake regulator" |
| CDS | /protein_id="AAK79675.1" /db_xref="GI:15024674" /translation="MTBKIPESDIEELSHSELIRMSWAEROYDCMEALEKODENMAE VTIKKDDIIDMOKEIENKTYIRLIMQOPVIAEBLNIPTVKIYVLDLELIGHANDI AKAIRLNGEKRHDIIVKINMKNKVSMTKDSLDATVERNLDAKAYVCRRDDYBSL YKRIFNELMLNTMSDBKRNQOLTQFLFCVLYLERIGRTTNVCSITLYLTGKQVDN D" |
| gene | 5586..6899 /gene="CACI710" 5586..6899 /gene="CACI710" /codon_start=1 /transl_table=11 /product="Fe-S oxidoreductase, related to NifB/MoaA family with PZD N-terminal domain" |
| CDS | /protein_id="AAK79676.1" /db_xref="GI:15024675" /translation="MENKTYKIEPOSIAEENGIEVGDLSISNGEKYKDIIDRFPLMA DEYVAVKIQTKNKIEIWEIIEKEVDEKLCVFBEFGIILDSAKSCRNKKIIFCFIDQLRKG |

MEKTLVFKDDSLRSLFLOGNFVLLNDDIDRI IKYRIS PINI SVQTNPELRKM
LNKXAGNIMEMRMKLSNAGITNACOVNCPGINSGEFSKTVLDLYKPSVNI
VPGITSYRBEFLNNPYDSQALKEINIKOEKFKESFVLSSEFVLSGI
PIPKAEYGEFEQEDGIMTRPNDINMLDKGHIKHSFMTVGLAVEIENA
AKMEXRSGIKIVIKVINEFKEKTVAGLIGRIILYKXNEVDYIILPSNL
KSDEVFLLDVKVLDIEKLNKKVLCVNSGEDIIVDINKNGRE"

gene

CDS

6902.8218
/gene="CAC1711"
/codon_start=1
/product="Predicted GTPase"
/protein_id="AAK79677.1"
/db_xref="GI:15024676"

gene

CDS

8220.9218
/gene="CAC1712"
/note="gpa"
/db_xref="GI:15024677"
/product="Glycerol 3-phosphate dehydrogenase"
/protein_id="AAK79678.1"
/db_xref="GI:15024677"

gene

CDS

9300.10778
/gene="CAC1713"
/db_xref="GI:15024678"
/product="Glycerol 3-phosphate dehydrogenase"
/protein_id="AAK79679.1"
/db_xref="GI:15024678"

Query Match 3.6%; Score 169.2; DB 1; Length 10861;

Best Local Similarity 50.5%; Pred. No. 1.1e-12;
Matches 498; Conservative 0; Mismatches 473; Indels 15; Gaps 3;

607 TTTCTGATATATAGTAAATAAATGAAATCAGTATTTAGGTGAGGATCATTTGG 666
8198 TTTAGGAAGAGAACTAATATAGGGGAGTAACTTACAGAGAGAGACCTTTGG 8257
667 CACAGCAATAGCAATTCGCTCAGCACTGATGATATCACTTAATGGGACGTGA 726
8258 TACAGCACTGCAATATGCTTCTAAAGAGTCAATATGATTTGGGATTAAGAA 8317
727 TCATGAAATATATACATATTAACACTTACCGAAAAATTTAAATTTACCCACATA 786
8318 TAAAGAAATTTCTGAAGATATTAATTAAGAACTAATACAGATATCTTCCAAACA 8377
787 TCATCTACAGCAACATATATGCAACAGCAATTAAGCAAGATATTTATGCAACAA 846
8378 CATTAATACCTGCTCGTTAAGCAGTATGATTTAGAGAG---CAGCTAAGGAAG 8434
847 TACATGATATATCTTAATCTTCAACAACTTACGCAATGATGATACAAATACA 906
8435 CAATATATATGTTCTTCTGCTGCCATCTTTTGCTTAAGGAGATGACAGAAATTA 8434
907 ACACAAACGACATATGTATAATCTCCATATTAATTTGTAGTAAGGATGCAAT 966
8495 GGGCTT-----TTTAAAGAGAGCAAAATATATTAGATTTGCAAAAGCAGTAGAGA 8548

QY 967 TACATCACTCAAAATTTCCAGTGAATAAGCAGAGAAATTTTACATATATCAATTTT 1026

DB 8549 AGAGACCAAGAGAGAGCTCTGGAAGTGTAAAGAGAACTTTAATAATCTGTTGT 8608

QY 1027 TATACCTCTGCTCCAGTTTGTCTAAAGAAATTTGACAGACCTCTTCTGTAGATGT 1086

DB 8609 AGTACTTCTGTGCTTACATGACAGAGAACTTCCAAATGATATTTCCACATCACTAGT 8668

QY 1087 ACTTGCTGTGATTAATTAAGAACTTGTGATCATCTGATAGAAACATTAAGATGATGT 1146

DB 8669 GGTACTTCTCAGAGATTAATTAAGAACTTGTGATAGATGATTTTATGATCAACAG 8728

QY 1147 TCTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1206

DB 8729 TTTTAGGCTGACACAAATGATGATTTAGCGCTTGAATAAGAGAGAGAGCTTAA 8788

QY 1207 CATTAATGCAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1266

DB 8789 TATATGCTCTGACATCTGCAATAGTATGATGATGATGATGATGATGATGATGAT 8848

QY 1267 TACTGTTAATTAAGCAGTAAATGAATTAACACTATATATAGCAAAATATCATT 1326

DB 8849 GGGCTTATGACAGAGAGTATGATGATGATGATGATGATGATGATGATGATGAT 8902

QY 1327 AATAGATCTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1386

DB 8903 AGGTAGCCAGAACTTTTGTGATTAAGCACTGATGATGATGATGATGATGATGATG 8962

QY 1387 AACAGAACATTCACGCAATATGCTTTTGGACTAGAAATAGGAAAGTAGAATAATA 1446

DB 8963 TAGTATGACAGTAAATTAAGAAAGCTGATGATGATGATGATGATGATGATGATG 9022

QY 1447 TACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1506

DB 9023 GGAAGCATGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 9082

QY 1507 ATCATTAAGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1566

DB 9083 TGATTAATTAAGAACTTGTGATGATGATGATGATGATGATGATGATGATGATG 9142

QY 1567 ACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1592

DB 9143 ATTTGAGAAATGAGATCTTAATAAAG 9168

RESULT 35

AY309266 3900 bp DNA linear BCT 22-SEP-2003

LOCUS AY309266

DEFINITION Borrelia burgdorferi N1pD (n1pD) and R1p2 (r1p2) genes, complete

VERSION AY309266

KEYWORDS AY309266.1 GI:33520757

SOURCE Borrelia burgdorferi (Lyme disease spirochete)

ORGANISM Borrelia burgdorferi

REFERENCE 1 (bases 1 to 3900)

AUTHORS Yang, X. F., Alani, S. M., and Norgard, M. V.

TITLE The response regulator R1p2 is essential for the expression of

JOURNAL major membrane lipoproteins in Borrelia burgdorferi (2003)

PROC. Natl. Acad. Sci. U.S.A. 100 (19), 11001-11006

REFERENCE 2 (bases 1 to 3900)

AUTHORS Yang, X. F., Norgard, M. V., and Alani, S. M.

TITLE Direct Submission

JOURNAL Submitted (29-MAY-2003) Microbiology, University of Texas

Southwestern Medical Center, 6000 Harry Hines, Dallas, TX 75390,

USA

FEATURES

source location/Qualifiers

1..3900

/organism="Borrelia burgdorferi"

```

/mol_type="genomic DNA"
/strain="297"
/db_xref="taxon:139"
223..1110
/gene="nlpd"
223..1110
/gene="nlpd"
/feature="nlpd"
/feature="putative membrane-bound metalloproteinase; similar
to ORF B80761 in Borrelia burgdorferi strain B31 genome"
/codon_start=1
/strand_table=11
/product="NLPD"
/protein_id="AA021217.1"
/db_xref="GI:33520758"
/translation="MSKIFLEFVAGFPEFLKIIYVSPYSEIKNSKODPVFSDIKIKVL
KYNKKOHIPLEFYKYKGGKDFEFLANKINGOSGATINILSDPAYSVEIYPS
KKGVPEDSKDYRNNLLATRDLTAEKVKIKENDRYVEFEFVKKPKDGLFSGT
ELIFLPIANRIFPLKKRIVISDRGPNDEFTGNKSPHNGIDLAAPMAEVSSSGIV
IEAGYNDIXNRYVGVHKNKIKSLYGHMLYSYKIGDPFKSEPFLEXGQGTGRATGPH
LHEFLKKNIPINPLKFLK"
complement(1359..2717)
/gene="rip2"
complement(1359..2717)
/feature="response regulator; enhancer-binding protein that
activates sigma 54-dependent genes; transcriptional
activator that controls expression of sigma 54"
/codon_start=1
/strand_table=11
/product="Rip2"
/protein_id="AA021218.1"
/db_xref="GI:33520759"
/translation="MSKIFVADDEKNIRKCIATYVLEDSYFVFTASDGEALETIENE
NLDVILSDRMPQISEKLIKIVKKNLIPILITLAVTSADVAMEGKYPDLTK
PDLERLLILIKSLINKKNDNDENLILIKDLKYEIMKSLMOYKIDPLVI
KIAKNSAVLITGESVGEKIIADAFDLSNRDQKPIVNCALASELSEBELGHE
KAGFTASISKKRPELANKGTFLDEIDELSEIQLKRVLOKNTFERRVGGATIK
VDRLIATNKRIEELIKKREKREDELFYRLININIPERKNDISYLNILIKDV
AKENNEEKLSDAMKALYYDPNGNIEIKNVLASALISKQGTIDEDIPATIKV
NENLIRKIDPISLEAKELIKQTLPHSKKSKCAEILIKGKILANKIIEYNI
DO"
ORIGIN
Query Match 3.64; Score 169; DB 1; Length 3900;
Best Local Similarity 49.54; Pred. No. 1,7e-12;
Matches 600; Conservative 0; Mismatches 575; Indels 36; Gaps 5;
QY 2661 TAAATGCGAATAGATCTATATCTAGTAAATCTTTGCTGATCATCTCCTTAG 2720
DB 1529 TAAATGCTTATATTTTATTTTGGTAAATCTT---CTTAGATTTGTTTGC 1585
QY 2721 GAGATTCATTAATAATGATCTATATACTAGTAAATGCTTAACTAGTACCGGCC 2780
DB 1586 CTTTGTATATTAATGCGCTTCAAGCATTTTAAATTTCTTAATATTTCTGAGCC 1645
QY 2781 ATTCAATGACTGATGCTATTAAAGCTTCACTAATAATGAGTACAAACCTTA 2840
DB 1646 AATCGATATATAGAGGCTTTCATGATCAATTAAGAAAGTTTCTCTCATATGT 1705
QY 2841 TTTTTCAGATGCTATTTTAAATACCTACACAAATCCGATATCTGACATATTT 2900
DB 1706 TTTTCCTTGGAGAGCTTTTATATGATGTTGTTAAATAGATATATCATCTTTCTT 1765
QY 2901 CTATTAAGAGGCTACTCTAATGAGAGCATTTATCTATATTAATAATCTCAGAAA 2960
DB 1766 CTCTTAAGGCGGATGTTATATGATATTTATCTATTAATAATATCTTCTTAA 1825
QY 2961 ACCTACAGCTTTTACTCTTCAATCTTGAAGAAAGACAAATATTTTCACT 3020
DB 1826 ATTTTCCTTTTATATTTCTCTCAATGTTTGTGTTGAGGCAAGAGCTGATAT 1885
QY 3021 CTATCTACAGAGATCTTATATTTTCTCATATATTTTCTCTCTGATTAATCTGA 3080

```

```

DB 1886 CAATTTATATGAGCTTCTCCCAACAGCTTCAAAAAGTTTGTGACGACTCTTA 1945
QY 3081 GTAATCTTAATGAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 3140
DB 1946 AAGGTTGACTGTAATTTGAGTAAATTTCTGATCTATCAAGAAAATTTGCTCT 2005
QY 3141 GATTTGCTTGTCTATATTTGATGATGATGATGATGATGATGATGATGATGATGAT 3200
DB 2006 TGTGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2065
QY 3201 TTTCTGCTACCAATATATATATATATATATATATATATATATATATATATATAT 3260
DB 2060 CTTTTCATGCGCAAAAAGTTCATTTCAAGAAATGCTTCAAGAAAGTCTGCGCAATTA 2119
QY 3261 TAGATATTAATGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3320
DB 2120 CTTTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2179
QY 3321 CTTTCCACACCTGCGGACCAAGTATGATGATGATGATGATGATGATGATGATGATGAT 3380
DB 2180 CTTTACCAACACCTGCGGACCAAGTATGATGATGATGATGATGATGATGATGATGATGAT 2239
QY 3381 TATCATGATCTTCTCAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3437
DB 2240 TTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2299
QY 3438 CAAATGCTGATTTCACTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3486
DB 2300 TTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2359
QY 3487 -----TATGCTCTCTTACAACTGATGATGATGATGATGATGATGATGATGATGATGAT 3539
DB 2360 TTTTATTTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2419
QY 3540 TATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3599
DB 2420 AAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2479
QY 3600 TATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3659
DB 2480 TTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2539
QY 3660 CATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3719
DB 2540 CAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2593
QY 3720 AAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3779
DB 2594 CAATGTTTCAAGGCTTCTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 2653
QY 3780 TTAATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3839
DB 2654 CAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2713
QY 3840 AAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3890
DB 2714 TCATTTTCAAT 2724

```

RESULT 36
 AC120883 133877 bp DNA linear HTG 06-AUG-2002
 LOCUS Homo sapiens chromosome 18 clone CTD-2146H24 map 18, WORKING DRAFT
 DEFINITION
 SEQUENCE, 4 ordered pieces.
 AC120883
 AC120883.2 GI:22123189
 VERSION HTG PHASES: HTGS_DRAFT; HTGS_FULLTOP.
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE 1 (bases 1 to 133877)
 Birten, B., Nusbaum, C. and Lander, E.

[illegible]

| | | | | |
|------------|--|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------|----------------------------|
| Db | | 95539 | ATATTGTTGGTATTAATTATATAATAAATAAAATATATTAATAATAATTAATAATAA | 95539 |
| Oy | | 1938 | TAACTAATAAAAAGAACCAATACATATTAATAATTCTAAAACCCTATTACTACGTAA | 1997 |
| Db | | 95559 | TGTAAATATAATATAAGTGGA--ATATAATATATATTTTATATATAATATATAATAT | 95655 |
| Oy | | 1998 | TATTAACAATAGTCCTCGACATATAAATTCATCATCCCTGCATATTTCCAACGTAA | 2057 |
| Db | | 95556 | TTATATATATAATATAATAATAATAATAATATAATATTTATTTATATAATATAATAA | 95715 |
| Oy | | 2058 | ATATGATAAAGTGTTTTTTTTANGCCATTAGTTCCTTAATATAATGCTAATACCAGAA | 2117 |
| Db | | 95716 | AATATAAATTAATTAATTTATTTATATATAATATAATAAATAATAATATAATTTATATA | 95775 |
| Oy | | 2118 | AGGTTAAATACACATTTCCCGCACATATGAGAAGATACAAATACAGCTCAATTAAT | 2177 |
| Db | | 95776 | ATATATATAAATAATAATAATATAATATATTTATATATAATATAATAATAATAATATTT | 95835 |
| Oy | | 2178 | AAACAAATTTCTAATATAACCAACCTACTTAAACATGGAACCAATAGTATTCATCCT | 2237 |
| Db | | 95836 | ATATAAATATAATAATAATAATATAATTTATATATAATATAATAATAATAATATTT | 95899 |
| Oy | | 2238 | AAGTAAGATATCCCTCAGACATGTATCAATAATAAAGAAAAATGCTTAGTATATTAT | 2297 |
| Db | | 95896 | ATATAATATAATAATAATAATAATAATATATTTATATATAATAATAATAATAATAAT | 95955 |
| Oy | | 2298 | ACAAATACAAATTTTCCAACTTCAACACAAACCAAACTCAAAGTCTAT-TATAACAAATCA | 2356 |
| Db | | 95956 | TTATATATAATATAATAATAATAATAATAATATATTTATATATAATAATAATAATAAT | 96015 |
| Oy | | 2357 | AGAAATATATATATCTGCCCAACACACATAGTATCAAAATTCATATATACATATCTTT | 2416 |
| Db | | 96016 | AATAATATTTATTTTATTTATATAATATATAATAATAATAATAATATATTTATTTATTAAT | 96075 |
| Oy | | 2417 | TTAAAGAATGACTTCGCAATTCATCATATAATAATAATAATACATGCAAGTTTATACAGAA | 2476 |
| Db | | 96076 | ATATAAATAATAATAATAATTTATTTATTTATATAATAATAATAATTTTAGTTATATATAAA | 96135 |
| Oy | | 2477 | ACATTACA 2484 | |
| Db | | 96136 | TATTTATA 96143 | |
| RESULT 37 | | | | |
| PFML4P3 | | | | |
| LOCUS | | | | |
| DEFINITION | | PFML4P3 | 349751 bp | DNA linear INV 29-JAN-2003 |
| ACCESSION | | AL035476 | AL844503 | |
| VERSION | | AL035476.7 | GI:23498195 | |
| KEYWORDS | | | | |
| SOURCE | | Plasmodium falciparum 3D7 | | |
| ORGANISM | | Plasmodium falciparum 3D7 | | |
| REFERENCE | | | | |
| AUTHORS | | 1 Hall,N., Pain,A., Berriman,M., Churcher,C., Harris,B., Harris,D., Mungall,K., Bowman,S., Atkin,R., Baker,S., Barron,A., Brooks,K., Buckee,C.O., Burrows,C., Cherevach,I., Chillingworth,T., Chillingworth,T., Christodoulou,Z., Clark,L., Clark,R., Corcoran,C., Cronin,A., Davies,R., Davis,P., Dear,P., Dearden,F., Doggett,J., Felwell,T.T., Goble,A., Goodhead,I., Gilliam,R., Hamlin,N., Hance,Z., Harper,D., Hauser,H., Hornsbay,T., Holroyd,S., Horrocks,P., Humphray,S., Jagels,K., James,K.D., Johnson,D., Kerhornou,A., Knights,A., Konfortov,C., Kyes,S., Larke,N., Lawson,D., Lennard,N., Line,A., Meddison,M., McLean,J., Mooney,P., Moulis,S., Murphy,L., Oliver,K., Ormond,D., Price,C., Quail,M.A., Rabbinowitsch,E., Rajandream,M.A., Rutter,S., Rutherford,K.M., Sanders,M., Simmonds,M., Seeger,K., Sharp,S., Smith,R., Squares,R., Squares,S., Stevens,K., Taylor,K., Tivey,A., Unwin,L., Whitchead,S., Woodward,J., Sulston,J.E., Craigs,A., Newbold,C. and Barrett,B.G. | | |
| TITLE | | Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13 | | |
| JOURNAL | | Nature 419 (6506), 527-531 (2002) | | |
| MEDLINE | | 2235708 | | |

PUBMED 12368867
2 (bases 1 to 349751)
REFERENCE Hamlin,N., Pain,A., Berriman,B., Hall,N., Bowman,S., Churcher,C.,
AUTHORS Harris,B., Harris,D., Lawson,D., Quail,M. and Barrett,B.
TITLE Direct Submission
JOURNAL Submitted (24-FEB-1999) P.falciparum Genome Sequencing Consortium,
The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge
CB10 1SA, UK
COMMENT On Oct 3, 2002 this sequence version replaced gi:5763808.
For more information about this sequence or the Malaria Project,
see http://www.sanger.ac.uk/Projects/P_falciparum.
FEATURES
source
1. 349751
/organism="Plasmodium falciparum 3D7"
/mol_type="genomic DNA"
/isolate="3D7"
/db_xref="taxon:36329"
/chromosome="4"
/clone="MAL4P3"
/complement(join(2130..3551,4472..9988))
/gene="PFD0630c"
/note="synonym: VAR"
/complement(join(2130..3551,4472..9988))
/note="similar to Plasmodium falciparum var. mal4p2.58
mal4p2.58, var SMALL:Q9U0G5 (EMBL:AL035475) (2277 aa)
fasta scores: E(): 1.5e-143, 52.77% id in 2486 aa, and to
Plasmodium falciparum variant-specific surface protein
var-7 SMALL:Q26034 (EMBL:L42636) (2182 aa) fasta scores:
E(): 1.9e-126, 46.72% id in 2412 aa"
/codon_start=1
/product="erythrocyte membrane protein 1 (PfEMP1)"
/protein_id="CAD49166.1"
/db_xref="GI:23498197"
/db_xref="GOA:Q811V7"
/cd_xref="SPTREMBL:Q811V7"
/translation="MGPPSTADPDSGSAKHALDSIGKRVHQAQVONAKORSNGDLKGL
TSATLSGGSAPFENPCILIKKDELLIGDSGRHPCNGSASERFSVSGGSDDK
KLEGNRNNGACAPYRRLFLCNKMGKSTTKHLLDVCMAAYEASLIRH
DOTDATYPSGDSFSCMTMLARSPADIGDITRGKDLVGRKKKKTETEDQLESKIKI
FGDINELNNGRNKGVNDHYQDNGGNGFQLEFREDWNTANRATVKAITCKADTAAYFR
PCTSNQGSQAHHYCRGNDGKPDPTDPTDYVQPLRWFEMWAFORCKK
IYGVIVTKYCRKYSNGENPRYCRNGYDCTKRAIGKRWNGCICLACAPYVY
MINNOECPDKOKKKTDEINBASRSRROKRGASTGSSSYVDGSEKFNILKDYV
GYVDFLKLINKEKSCQAVKNDGGTINTEKNDKNNNNKNGKGFYSFYSCVCPDC
KYNVSGWEEKKNDCKIKYRKKAAPHRITIKLSGKEEIEKRLAFDCKD
GKNSDSLYDWCQCFDQLEKDEKEBGVDNRYNDVTRGGLCTIDKRGSENGKK
QTYNDFNFWAHMLKDSIHWKKLKELQNGITRCRNNKCKECCORWVEKK
KTEWGIKDFXTONIGDETNDPTVLEGVAKQFLNEDSYQDKONLDBEELHKK
HSEMLOETSGDLTCGASDNKETIMDLIHEBGIAECLAKQNECEKAKEGRS
DSHDQPPDDADNEDLDDEDEDEEVEVDNTOEGEGVYVQCEGSSSPTPA
PAGPDYDVISQLEFSDPSQFACITLKYTGKNGMKICPSGNTDITGSSSPATGAR
GRDDTSGDITGKDGATGCIIPRRRLIYVGLISOMASRTOGTGGSSGNILAE
AIIQSAIITFFLMKRYKAKENTTQVAGAGADPAPASSVATALAPAVSRSLQ
LSGVGVPBPMAVYKSIIPPLGVGVGIPVAGLGGAGLGGPGLGVGQAO
PILRLPVLGNGLOSPOGRRLTLGHPAGSGSEDTPOWLOQGIIPDLRLMFTY
LADYDICIGGRDVIAGTIVNTSGSSSKIKISEKIELHNNKPEAPKSVET
TPTWMSNGOHIWNGMICALTYDSSGISTEDQVADKILIDKDTGPOKNGNDYV
YERVELKDESGPKNDITOPATIKDFEILPYFRMLHEMGSDFGKAGKIKDYKON
CENIDKAGHHYSGGQYDCTPRVTERNKDFYDNLGCKYKOKRKKKIDIKPYEYH
QEKTKDEYGLTKDSDDKLEGRKCHENFLKELKRCRSEDNNDDNKINPDKPE
KTFNPKYCRKAPVGVKTYGNYIPNEKDKSKRGRVKEENDTIPNILEVLVGR
GEKDKDHLHACKNTGLFEVARYEOWNQCKKGIIDOKITKFPANDIDFPDILVFN
FQPMRLYEVQDYNKDKINCTKETEKEKEKYSKTOGCDNKECEKEMLKQ
EMTQIKTLKYOKYSIDOBIAFRKYSFVQDQIPNDYKAKOEVKECDKELMGCT
GNILKGEDEGKCMGDFITNLISLQKDKIDCNQAOANSVETPOSEPNACODDT
PDEEDLLEENNOVAOBNICNOVEDKIEEVEKCEIATQTAETAAAGSGRQTP
PAPAPAPSPRPRLPKRPPRPDLPALKAMLSSTIMSYIGRAAFYFPIKKT
KASVGNLFILOPKGDYDITPKSSNRTIPASRYGKATTIMEGSSDEKAEAM
SOTDVTSSSEEBELDVNDIYVPSPKTKILEVLVPSGNTPTESIPDNTPTQ
PITDDEMNQDKHFIISMLQNTQNTENPILHNDVNTHPITMSRMDQKFIISID
RLNFSGEENYMFNSGNPILNIDSTNSMDLISNHSYNDKIDLDSGLDINDAL
SGNHIDYDEMLKRENELFGTHQHPKNTSNRVVVTQTSDDPPIHNOQLPHTMLDRH

misc_feature
13697..13898
/note="GC-rich region"
/complement(join(17269..18690,19611..25169))
/gene="PFD0635c"
/note="synonym: VAR"
/complement(join(17269..18690,19611..25169))
/note="similar to Plasmodium falciparum var. mal4p2.58
mal4p2.58, var SMALL:Q9U0G5 (EMBL:AL035475) (2277 aa)
fasta scores: E(): 1.1e-129, 49.19% id in 2498 aa, and to
Plasmodium falciparum variant-specific surface protein
var-7 SMALL:Q26034 (EMBL:L42636) (2182 aa) fasta scores:
E(): 7.9e-126, 46.25% id in 2441 aa"
/codon_start=1
/product="erythrocyte membrane protein 1 (PfEMP1)"
/protein_id="CAD49167.1"
/db_xref="GI:23498197"
/db_xref="GOA:Q811V6"
/cd_xref="SPTREMBL:Q811V6"
/translation="MAPTSGGGETKDSAKHALDSIGKRVHQAQVONAKORSNGDLKGL
LTSATLSGGSAPFENPCILIKKDELLIGDSGRHPCNGSASERFSVSGGSDDK
KLEGNRNNGACAPYRRLFLCNKMGKSTTKHLLDVCMAAYEASLIRH
DOTDATYPSGDSFSCMTMLARSPADIGDITRGKDLVGRKKKKTETEDQLESKIKI
FGDINELNNGRNKGVNDHYQDNGGNGFQLEFREDWNTANRATVKAITCKADTAAYFR
PCTSNQGSQAHHYCRGNDGKPDPTDPTDYVQPLRWFEMWAFORCKK
IYGVIVTKYCRKYSNGENPRYCRNGYDCTKRAIGKRWNGCICLACAPYVY
MINNOECPDKOKKKTDEINBASRSRROKRGASTGSSSYVDGSEKFNILKDYV
GYVDFLKLINKEKSCQAVKNDGGTINTEKNDKNNNNKNGKGFYSFYSCVCPDC
KYNVSGWEEKKNDCKIKYRKKAAPHRITIKLSGKEEIEKRLAFDCKD
GKNSDSLYDWCQCFDQLEKDEKEBGVDNRYNDVTRGGLCTIDKRGSENGKK
QTYNDFNFWAHMLKDSIHWKKLKELQNGITRCRNNKCKECCORWVEKK
KTEWGIKDFXTONIGDETNDPTVLEGVAKQFLNEDSYQDKONLDBEELHKK
HSEMLOETSGDLTCGASDNKETIMDLIHEBGIAECLAKQNECEKAKEGRS
DSHDQPPDDADNEDLDDEDEDEEVEVDNTOEGEGVYVQCEGSSSPTPA
PAGPDYDVISQLEFSDPSQFACITLKYTGKNGMKICPSGNTDITGSSSPATGAR
GRDDTSGDITGKDGATGCIIPRRRLIYVGLISOMASRTOGTGGSSGNILAE
AIIQSAIITFFLMKRYKAKENTTQVAGAGADPAPASSVATALAPAVSRSLQ
LSGVGVPBPMAVYKSIIPPLGVGVGIPVAGLGGAGLGGPGLGVGQAO
PILRLPVLGNGLOSPOGRRLTLGHPAGSGSEDTPOWLOQGIIPDLRLMFTY
LADYDICIGGRDVIAGTIVNTSGSSSKIKISEKIELHNNKPEAPKSVET
TPTWMSNGOHIWNGMICALTYDSSGISTEDQVADKILIDKDTGPOKNGNDYV
YERVELKDESGPKNDITOPATIKDFEILPYFRMLHEMGSDFGKAGKIKDYKON
CENIDKAGHHYSGGQYDCTPRVTERNKDFYDNLGCKYKOKRKKKIDIKPYEYH
QEKTKDEYGLTKDSDDKLEGRKCHENFLKELKRCRSEDNNDDNKINPDKPE
KTFNPKYCRKAPVGVKTYGNYIPNEKDKSKRGRVKEENDTIPNILEVLVGR
GEKDKDHLHACKNTGLFEVARYEOWNQCKKGIIDOKITKFPANDIDFPDILVFN
FQPMRLYEVQDYNKDKINCTKETEKEKEKYSKTOGCDNKECEKEMLKQ
EMTQIKTLKYOKYSIDOBIAFRKYSFVQDQIPNDYKAKOEVKECDKELMGCT
GNILKGEDEGKCMGDFITNLISLQKDKIDCNQAOANSVETPOSEPNACODDT
PDEEDLLEENNOVAOBNICNOVEDKIEEVEKCEIATQTAETAAAGSGRQTP
PAPAPAPSPRPRLPKRPPRPDLPALKAMLSSTIMSYIGRAAFYFPIKKT
KASVGNLFILOPKGDYDITPKSSNRTIPASRYGKATTIMEGSSDEKAEAM
SOTDVTSSSEEBELDVNDIYVPSPKTKILEVLVPSGNTPTESIPDNTPTQ
PITDDEMNQDKHFIISMLQNTQNTENPILHNDVNTHPITMSRMDQKFIISID
RLNFSGEENYMFNSGNPILNIDSTNSMDLISNHSYNDKIDLDSGLDINDAL
SGNHIDYDEMLKRENELFGTHQHPKNTSNRVVVTQTSDDPPIHNOQLPHTMLDRH

gene
/note="synonym: RIF"
/complement(join(28486..29547,29708..29776))
/gene="PFD0640c"
/note="synonym: RIF"
/complement(join(28486..29547,29708..29776))
/note="similar to Plasmodium falciparum rifin pfel115w,
mal3p.54 SMALL:O97311 (EMBL:AL034559) (371 aa) fasta
scores: E(): 4.8e-58, 49.73% id in 382 aa, and to
Plasmodium falciparum rifin pfb0035c SMALL:O96113

QY 1980 CACCTATTACTAGTATATTCATAGATGTCCTGACATATTAATCTCATCTCTG 2039
 Db AATTTATTTGAAAATATATATATATATATATATATATATATATATATATATAT 68805
 QY 2040 ACATATTTCAACATATAATATATATATATATATATATATATATATATATAT 2099
 Db AATATATATATATATATATATATATATATATATATATATATATATATATATAT 68806
 QY 2100 ATTGCTAATTTACCAAAAAGTTAAATATACATCTCCGACCAATATGAGAGATCA 2159
 Db AATATATATATATATATATATATATATATATATATATATATATATATATATAT 68866
 QY 2160 ATACAGCTCAAAATATATATATATATATATATATATATATATATATATATAT 2219
 Db AATTTATATATATATATATATATATATATATATATATATATATATATATATAT 68926
 QY 2220 CAATAGTATATATATATATATATATATATATATATATATATATATATATAT 2279
 Db TATATATATATATATATATATATATATATATATATATATATATATATATATAT 68985
 QY 2280 ATGCTATATATATATATATATATATATATATATATATATATATATATATAT 2339
 Db TATATATATATATATATATATATATATATATATATATATATATATATATATAT 69045
 QY 2340 CTATATTAACAAATTCAGATATATATATATATATATATATATATATATATAT 2398
 Db AATTTATATATATATATATATATATATATATATATATATATATATATATATAT 69105
 QY 2399 TATATATATATATATATATATATATATATATATATATATATATATATATAT 2458
 Db TATATATATATATATATATATATATATATATATATATATATATATATATATAT 69165
 QY 2459 TGCATATATATATATATATATATATATATATATATATATATATATATATAT 2510
 Db TATATATATATATATATATATATATATATATATATATATATATATATATATAT 69225
 QY 2511 -----GTATGCTGACGCTTCATTCACAAATTCAGAGTGTGATACATTA 2564
 Db AATATATATATATATATATATATATATATATATATATATATATATATATATAT 69285
 QY 2565 CAAAACGATATATATATATATATATATATATATATATATATATATATATATAT 2624
 Db TATATATATATATATATATATATATATATATATATATATATATATATATATAT 69345
 QY 2625 ATGGTACAGAAATATATATATATATATATATATATATATATATATATATAT 2684
 Db TATATATATATATATATATATATATATATATATATATATATATATATATATAT 69405
 QY 2685 CTACGTGTAATCTTTTGTGCTATCATCTCCCTGAGAGATTCATATATATATAT 2744
 Db ATATATATATATATATATATATATATATATATATATATATATATATATATAT 69465
 QY 2745 CTATATATATATATATATATATATATATATATATATATATATATATATAT 2804
 Db TGAATATATATATATATATATATATATATATATATATATATATATATATATAT 69525
 QY 2805 AAGCTCATCATCTATATATATATATATATATATATATATATATATATATATAT 2858
 Db TTAATATATATATATATATATATATATATATATATATATATATATATATATAT 69585
 QY 2859 TCATATATATATATATATATATATATATATATATATATATATATATATATAT 2918
 Db TATATATATATATATATATATATATATATATATATATATATATATATATATAT 69645
 QY 2919 TATATGAGAGACATATATATATATATATATATATATATATATATATATATAT 2978
 Db AATATATATATATATATATATATATATATATATATATATATATATATATATAT 69705
 QY 2979 CACTTCAATATCTTTGAGAGAGACATATATATATATATATATATATATATATAT 3038
 Db AATATATATATATATATATATATATATATATATATATATATATATATATATAT 69765

QY 3039 TACTATTTTCCCTATATATATATATATATATATATATATATATATATATATAT 3098
 Db TATATATATATATATATATATATATATATATATATATATATATATATATATAT 69825
 QY 3099 CGTATGCTAATCTGTATCTATATATATATATATATATATATATATATATATAT 3158
 Db ATATATATATATATATATATATATATATATATATATATATATATATATATATAT 69885
 QY 3159 TTCCATATGAGAGAGTATCTATATATATATATATATATATATATATATATATAT 3218
 Db TACATGATATATATATATATATATATATATATATATATATATATATATATATAT 69944
 QY 3219 TATTAACCAAGTATAT 3234
 Db TTTATATATATATATAT 70019
 RESULT 38
 AX622666 996 bp DNA linear PAT 20-FEB-2003
 LOCUS AX622666
 DEFINITION Sequence 5629 from Patent WO02094868.
 ACCESSION AX622666
 VERSION AX622666.1 GI:28450652
 KEYWORDS
 SOURCE Staphylococcus aureus
 ORGANISM Staphylococcus aureus
 Bacteria; Firmicutes; Bacillales; Staphylococcus.
 REFERENCE
 1. Mesigian, V.C., Mora, M.C. and Scarselli, M.C.
 Staphylococcus aureus proteins and nucleic acids
 Patent: WO 02094868-A 5629 28-NOV-2002.
 JOURNAL Chiron Spa (IT)
 FEATURES
 source
 1. 996
 /organism="Staphylococcus aureus"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1280"
 ORIGIN
 Query Match 3.5%; Score 168.6; DB 6; Length 996;
 Best Local Similarity 51.4%; Pred. No. 3.1e-12;
 Matches 505; Conservative 0; Mismatches 454; Indels 24; Gaps 4;
 QY 635 AATATCGATTTTATGAGTATGATATATATATATATATATATATATATATATATAT 694
 Db AATATCGATTTTATGAGTATGATATATATATATATATATATATATATATATATATAT 7
 QY 695 CATGATATATATATATATATATATATATATATATATATATATATATATATATATAT 754
 Db AATGATATATATATATATATATATATATATATATATATATATATATATATATATAT 67
 QY 755 TACGGAATATATATATATATATATATATATATATATATATATATATATATATATAT 814
 Db TGTATATATATATATATATATATATATATATATATATATATATATATATATATAT 127
 QY 815 AGCAATATATATATATATATATATATATATATATATATATATATATATATATATAT 874
 Db TCAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 187
 QY 875 CAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 934
 Db AAGCAATATATATATATATATATATATATATATATATATATATATATATATATATAT 244
 QY 935 CCAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 994
 Db ACTTATATATATATATATATATATATATATATATATATATATATATATATATATATAT 298
 QY 995 GCAGAAATATATATATATATATATATATATATATATATATATATATATATATATAT 1045
 Db ATGAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 358
 QY 1046 TTGCTAAGAAATGAGAAATCTTCTGTATATATATATATATATATATATATATATAT 1105

| Db | Accession | LOCUS | DEFINITION | VERSION | KEYWORDS | SOURCE | ORGANISM | REFERENCE | AUTHORS |
|------------|-----------|-------------------------------------------------------------------|----------------------------------------------------------------------------|---------|-----------------|--------|----------|-----------|---------|
| Db | 418 | CATGGGAGGAAGAGTTGTAGTACGACCAACCACTACAGTCTGCTGCTCATCAACAAAATATAA | | | | | | | |
| QY | 1106 | GAACCTGGTGAAATCATCTTGATGATGAAACAATAAGATGATGCTCTAAAAATAATATACAT | | | | | | | |
| Db | 478 | AGGTGATAGTAATTAATTAACCCAGATTTATTATGAAATGATTAATTTGGTGCTGACACGAT | | | | | | | |
| QY | 1166 | CAGAGATATATATAGGTGATGACATTTGGAGCTGCATTAATGAACCAATATTCATTCGATCT | | | | | | | |
| Db | 538 | GATGACTTATTTGGTGCTTGAACCTGGTGCTGATTAATAAATATCATGGCACTACCAAGT | | | | | | | |
| QY | 1226 | GGATATATGCTGTGAAAAAATTTTAGGTATATATGCTGTTGCTACTGTTTAACTTAAAGC | | | | | | | |
| Db | 598 | GGATATGCTGATCGGAATTTGGCTAGCGGTATATATGCAAAAGCTGCATTAATGATCTGTGC | | | | | | | |
| QY | 1286 | ATCAATAGTAATTAATAACCTATATATATGCAAAAAATCATTCATATAGATCTTCATCATTTA | | | | | | | |
| Db | 658 | TTTGCGGAAATTAATGAGA-----TTAGTGGAAAAAGTTAGTGCCGATCTCTATGACATTT | | | | | | | |
| QY | 1346 | ATTGTTCATCATCTGCTTTGGAGATCTTAATATTAACATGATCAACAGACATTCAGCCAT | | | | | | | |
| Db | 712 | CTAGGTTTATGGTGATGCTGATCGGTGACTTAATCGTTCTTGCACTCAACACATATCTCGAAT | | | | | | | |
| QY | 1406 | ATGCGCTTTGGAGCTGAAATATGAAAAAGGTGTGAATTAATATCATTAATATATGATGCAC | | | | | | | |
| Db | 772 | TTACACTTATGATATTAACCTTGACACAGATTAATCAATGATATACAGATTTATCTGAATG | | | | | | | |
| QY | 1466 | CTAAAGCTTTGTGAAGAACCGATCTGTAAACCACTGATATATCATTTAGCAAAAAAATT | | | | | | | |
| Db | 832 | AATATGTTGTTGGAAGTATTTATATCAACTAAATCAAGTATATCATTTAGCTAAAGAAAA | | | | | | | |
| QY | 1526 | AATGTGAATCTCCCAATTTTGCAATATCTTTACAAATTTATACATGAGAAATATATACTA | | | | | | | |
| Db | 892 | AATGTGATATGCAATTAATTCAAATGCAATTAATATATGAGTATTAATTTGAAAAATATCTCAGT | | | | | | | |
| QY | 1586 | GATTAAGCCATATCCAAACATATT 1608 | | | | | | | |
| Db | 952 | AAGATATGCTTAAAGATTAT 974 | | | | | | | |
| RESULT 39 | | | | | | | | | |
| PFMALIP3 | | 67970 bp | DNA | linear | INV 29-JUN-2003 | | | | |
| LOCUS | | | | | | | | | |
| DEFINITION | | | Plasmodium falciplarum DNA from MLIP3. | | | | | | |
| ACCESSION | | | AL031746 AL844501 | | | | | | |
| VERSION | | | AL031746.9 GI:6594243 | | | | | | |
| KEYWORDS | | | | | | | | | |
| SOURCE | | | Plasmodium falciplarum 3D7 | | | | | | |
| ORGANISM | | | Plasmodium falciplarum | | | | | | |
| REFERENCE | | | Enkaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium. | | | | | | |
| AUTHORS | | | 1. Hall, N., Pain, A., Berriman, M., Churcher, C., Harris, B., Harris, D., | | | | | | |
| | | | Mungall, K., Bowman, S., Atkin, R., Baker, S., Barron, A., Brooks, K., | | | | | | |
| | | | Bucke, C.O., Burrows, C., Cheruvach, I., Chillingworth, C., | | | | | | |
| | | | Chillingworth, T., Christodoulou Z., Clark, L., Clark, R., Cotton, C., | | | | | | |
| | | | Cronin, A., Davies, R., Davis, P., Dear, P., Dearden, F., Doggett, J., | | | | | | |
| | | | Feltriwell, T., Goble, A., Goodhead, I., Gwilliam, R., Hamlin, N., | | | | | | |
| | | | Hance, Z., Harper, D., Hauser, H., Hornsby, T., Holyoak, S., | | | | | | |
| | | | Horrocks, P., Humphray, S., Jagels, K., James, K.D., Johnson, D., | | | | | | |
| | | | Kochornou, A., Knights, A., Konfortov, B., Kyes, S., Laake, N., | | | | | | |
| | | | Lawson, D., Lennard, N., Line, A., Maddison, M., McLean, J., Mooney, P., | | | | | | |
| | | | Moule, S., Murphy, L., Oliver, K., Ormond, D., Price, C., Quail, M.A., | | | | | | |
| | | | Rabbinowitch, E., Rajandream, M.A., Rutter, S., Rutherford, K.M., | | | | | | |
| | | | Sanders, M., Simmonds, M., Seeger, K., Sharp, S., Smith, R., Squares, R., | | | | | | |
| | | | Squires, S., Stevens, K., Taylor, K., Tivey, A., Unwin, J., | | | | | | |
| | | | Whithead, S., Woodward, J., Sulston, D.E., Craig, A., Newbold, C. and | | | | | | |
| | | | Bartell, B.G. | | | | | | |
| TITLE | | | Sequence of Plasmodium falciplarum chromosomes 1, 3-9 and 13 | | | | | | |
| JOURNAL | | | Nature 419 (6906), 527-531 (2002) | | | | | | |

TITLE Direct Submission
 JOURNAL Submitted (24-SEP-1998) P.falciparum Genome Sequencing Consortium, The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK
 COMMENT On Dec 16, 1999 this sequence version replaced gi:5763807. For more information about this sequence or the Malaria Project, see http://www.sanger.ac.uk/Projects/P_falciparum.
 FEATURES
 source
 1..67970
 /organism="Plasmodium falciparum 3D7"
 /mol_type="genomic DNA"
 /isolate="3D7"
 /db_xref="taxon:36329"
 /chromosome="1"
 complement(join(201..257,455..586,748..804))
 /genes="PFA0575C"
 /note="synonym: MAL1P3.01b"
 complement(join(201..257,455..586,748..804))
 /genes="PFA0575C"
 /note="Similar to Schizosaccharomyces pombe hypothetical protein c119.18 spbcl19.18 SMALL:Q96VG1 (EMBL:AL022117) (69 aa) fasta scores: E(): 1.3e-06, 39.68% id in 63 aa, and to Homo sapiens protein 15el.1.15el.1.1 SMALL:15EL1 HUMAN (SWALL:O43715) (76 aa) fasta scores: E(): 0.00046, 38% id in 50 aa, and to Mus musculus protein 15el.1.15el.1.1 SMALL:Q9D622 (EMBL:AK007514) (76 aa) fasta scores: E(): 0.00054, 38% id in 50 aa"
 /codon_start=1
 /product="hypothetical protein, conserved"
 /protein_id="CAD49075.1"
 /db_xref="GI:23477009"
 /db_xref="SPTREMBL:O81227"
 /translation="MENRSKTIQNTKAHVEANDECKEKEXKLCFFNNWYKNNPLKGDILQADDYEDIQICLVNLDNKKGIIGHLSNVEKE"
 complement(join(1392..1461,1752..2598,2748..2848,2990..3276))
 /genes="PFA0580C"
 /note="synonym: MAL1P3.01"
 complement(join(1392..1461,1752..2598,2748..2848,2990..3276))
 /genes="PFA0580C"
 /note="Originally annotated as Plasmodium falciparum; conserved hypothetical protein but Pfam match and other evidences suggest that it is a putative deoxyribonuclease; earlier start site possible to make a 412 aa protein Pfam match to entry PF01026 Tard DNase, Tard related DNase, score 117.30, E-value 2.9e-31
 Similar to Plasmodium falciparum conserved hypothetical protein, upf0006 family mal1p3.01 SMALL:O9DUN6 (EMBL:AL031746) (412 aa) fasta scores: E(): 3.7e-150, 100% id in 412 aa, and to Saccharomyces cerevisiae putative deoxyribonuclease yb1055c yb1055c or yb10512 or yb10511 SMALL:YBFS YERAS7 (SWALL:P34220) (418 aa) fasta scores: E(): 2.5e-19, 30.47% id in 397 aa"
 /codon_start=1
 /product="Tard-like deoxyribonuclease, putative"
 /protein_id="CAD49076.1"
 /db_xref="GI:23477010"
 /db_xref="SPTREMBL:O81226"
 /translation="MKLVFVYKISYNLFIYSILFLKNSLKIYNDRYISTVANKRYKLOIKRGNLKNKNEIRKMDENESFIDIGNLDYKMDGVNASKHENDILQNVNAPKAMNNVDKLIITCTGLAEIDSKLICEITDPEGPKFYLSAGVATCYCFIDNKKEEKELI1AKKEEETIKYKROVNSKEMENKKIKCGEKMMNNLNELENDITIGPKFKIYKQKIVNNEKDXYENLKNKLTIKYPRIVQIGIGLDPRIYKCSGYIOIKYKIFOLKIYQMVNQLPFWLHNRKSETFKIVDYKFLPEKNGVHSFTFKEDIVHIVNKNYKLYLGGNGCSLSEINAVAKKIPILNLLISTDAWCGVQKTHASVYIKDYKRAIYNLKIKKIKI1KCDNTIKFERNEPNIADIATITKYKAEAAVFDLPCKK"
 5005..5496
 /genes="MAL1P3.02"
 /note="synonym: PFA0585W"
 5005..5496
 /genes="MAL1P3.02"
 5005..5496

```
/note="contains possible signal sequence. Signal peptide
predicted SignalP 2.0 HMM (Signal peptide probability
0.640, signal anchor probability 0.357) with cleavage site
probability 0.504 between residues 27 and 28. Scansegp hit
to P800867, Carbamoyl-phosphate synthase subdomain
signature 2"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAB63557.1"
/db_xref="GI:6594245"
/db_xref="GOA:G9U0N5"
/db_xref="SPRMBL:G9U0N5"
/translators="MKLNNRPPVLCPIILIFPLNSVYLGNNRRNNINFEENAK
ARKRLSGEINSIKLNDGDELKIKLNDKHKSTWDYSFISLSEKISQTLPR
KQOEINEANTKIIDRQEFYILNDEINIAIFVLENNFELYIQSPKSLIDIIOS
LNN"
/misc_feature
8179..10303
/note="Putative centromere"
14884..20352
/gene="MAL1P3.03"
/note="synonym: PFA0590W"
14884..20352
/gene="MAL1P3.03"
/note="Putative ABC transporter
pfam match to entry PF00005 ABC tran, ABC transporter,
score 45.00, E-value 1.7e-09, HMMstart hit to SMO0382,
Atpases associated with a variety of cellular activities;
ScanRegexp hit to P500211, ABC transporters family
signature
Similar to Schizosaccharomyces pombe ATP-binding cassette
transporter abci1 or spac99.12C SMALL:ABCI1 SCPO
(SMALL:092337) (1427 aa) fasta scores: E(): 4.2e-08,
21.468 id in 1146 aa"
/codon_start=1
/product="ABC transporter, putative"
/protein_id="CAB63558.1"
/db_xref="GI:6594246"
/db_xref="GOA:G9U0N4"
/db_xref="SPRMBL:G9U0N4"
/translators="MTTYKENYGISNKKKKKSCONTSPINLSFDMIRPLINDLIK
GDIQLPNTGRNDVPIYASKLEBNRDRDVEDSEFSKSNENHYAHKCHSDASE
KQVNVYHNLISILIKTRPILITISYILETLIVTGKGFIDYIMELISGQILPV
YISFLKDFVSGLVVVMIMFHLFEALHAFYFHLFTINLVSIMYLYIKINCSN
NHLNDAPANTYRKFSQTEIDELSRDLISGNASSSSGQIKNNNNINDKKEVEN
DYIINIKSKTKMEKDSLNENSLPNNYINIMFSDVPEVTFVSCINLEFVKIF
MSYVFIKIGNSVGIAIWMISALYSAMILPEPLPSLPSKYLTYDRKIDNMHVL
KPKLIDPMWSEFASKYINIPRMKMKYCKIRIYINICVPSISISDIDVVFYFI
YKDRLINKBEIKFTSIIMPLVYKLLISNVANPRLVNNWEGIVNKRKNYINDE
LYNDIKNFYRTRYNEDYINVDKTLQNNENITSHDGIHSHLHKNVYKQLIN
MPEYFFYHNSNNHKNIIKQILSGLKNDNTWKICFOEHSKSNVYNNSSIHE
KKEEYENIHNSNSTWSNEFEKKNNEYIKLENCSPGLSDKNCDDHILKINEN
LKNSMLIIGVSGSKAFSILGDPMTGNUNYIEMFCKMILVYPOQSWLFCG
NIRSLIFGEVNYPLKYKYLIOSELNDLSTIEGDMYINDNRILSKGQVLCIA
BALYEHYHMKLCTDYERKLIQNNELIKDLINKNISYNNKSKLVANNPENEN
YLOKCLMDNNKPLYLILDLITSIDPSISIKTSFNKICEDNISKDNSSPILNKS
LNDNFILEDLIDVQYEVNIFELQDKTLTKRGNISYMKNNLNTTKSHWYSUANT
IDYTRIKLDEYEAHVKRSHKNMAYEAFVKNESVSFEIDSINKETKCKKRY
KKEHNNKNNKNNNNNNNSKDDHININNDHNRNNDINDLPNSDTDSPLVSGNE
YLDUTYSNNSDKEEIVKPLKDTHEEFSKSMPEVSKSNINNSKFKEDSSS
FKGSISETLYAFQOVGFYLLTSVIFMLISFIDEIFVLTMMSSISKKNKSHD
TILOKOYLYEVFVILPISLTSVSGICFEMITTYGNTSAILKNNHILSLAPLYIF
YNNNIGNIIRREPIISADPYGLKRIYAPFPPFCILSILITIMYRDCIFEPV
IILYFVFRBSRCKEAKQRYLSCHTPLCNISYALSCKNITINIKYNTYHLYVE
HYNNRFRISFFRMLINIMASLYIKIFILITVYIMHSHLASGILIKYKKNRYI
LSTLGCISFSARLGVIIKFLCDYTHIEKMCVGRLEBPKISNKAASWKNEL
NVITTOYVEKENNISDKLSAIVEYKNSVSIINSODSEKRYGILFENVVYSY
KKIPLVNGYKYIDEEPSEKINMVALKNQIGIVGSDAGSKTILSLISGKVRICIA
KITVEGDLITVRKEDSDIIGLAQSFYVNNIRITIDPNNNTDEIVHALKN
GILKRGDLKYVHKODMSNYKIIOTKVFVNNOSNDNTIILINDCIRISVLYLN
RHXYKILIDELIPIPLNLSVHDEINSPLIGRAKSPNYIRRHFPNNVYLISHHANT
LSCDDIYVLRKKEITVRCSEYDEVKIQSELSHLEND"
/rRNA
23896..26044
/product="18S ribosomal RNA A-type"
```

```
misc_RNA
26045..26438
/product="18S A-type"
/note="ribosomal operon internal transcribed spacer 1
expressed in asexual stage parasites"
26439..26559
/gene="mal1.5.8rRNA"
26439..26559
/gene="mal1.5.8rRNA"
/note="5.8S ribosomal RNA"
/note="5.8S ribosomal expressed in asexual stage
parasites
hit to vivax A-type 5.8S ribosomal percent id: 88.00 none"
26560..27435
/product="18S A-type"
/note="ribosomal operon internal transcribed spacer 2
expressed in asexual stage parasites"
27436..31539
/product="28S ribosomal RNA (A-type)"
/note="28S rRNA expressed during asexual development"
complement(join(31566..32476,32675..32775))
/gene="MAL1P3.04"
gene
rRNA
misc_RNA
Query Match 3.5%; Score 168.4; DB 3; Length 67970;
Best Local Similarity 43.8%; Pred. No. 7.2e-13;
Matches 1227; Conservative 0; Mismatches 1541; Indels 34; Gaps 10;
328 ATTTCGAATTTATTTACAAATCATGAGCTCTTATCGACATACATTAATCAAT 387
7657 ATTATTTATTTACTTACACCGATATAGAAATAGTATTTCAATTAATCAAT 7716
388 CAATACAAATTAATAGTACTATCTAATTAATAAAAAATCCCATATAGTCTTACAGA 447
7717 TATATGATATATATTTCTTATTAATGATGCT-----CTCATTAATAGTACTTATGAT 7770
448 ACTACAAATTAATAGACACTTAATCTTATTAATATTTCAATTTATTAATCTCA 507
7771 TCTAAATATTTATTTGAATTAAGATTAATTAATTTCTAAGAAAGGACTTTATTAATA 7830
508 TCATATTAGCAATATTCATCTGACATGATTAACAATAGTAAACATTTCTTACAC 567
7831 ATATTCCTTATTTATTAATCAATTTAATAAATATTAATTAAGAAATTAATTTCTTA 7890
568 TAGATAAT-ATACATAAATTAATATATATTTTACTGCTGTTCTTGATTAATAGT 626
7891 AAACGATGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7950
627 AAAAAATGAATAGATTTTATAGGAGAGATCATTTGGACAGCAATAGCAATGGAC 686
7951 ATTTCATTAATTTTATTAATGTTTATTAATTAATTAATTAATTTCTTATTAATTAAT 8010
687 TGTGACAGCAATGTATATCAATTAATTAATTAATTAATTAATTAATTAATTAAT 746
8011 TATAGTACACTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTT 8060
747 TAAACACTTACCGAAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 806
8061 TGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8120
807 ATGCAACCGACATATAGACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 866
8121 ATATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8180
867 TTCTCAACAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 926
8181 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8240
927 AAAATACTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 986
8241 AAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8300
987 GTGAATATGAGAGAAATTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1046
```

[illegible]

DEFINITION Drosophila melanogaster complete mitochondrial genome.
 ACCESSION U37541
 VERSION U37541.1 GI:1166529
 KEYWORDS Mitochondrion Drosophila melanogaster (fruit fly)
 SOURCE Drosophila melanogaster
 ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Preygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 REFERENCE 1 (bases 12511 to 12682)
 AUTHORS Clary,D.O., Goddard,J.M., Martin,S.C., Fauron,C.M. and Wolstenholme,D.R.
 TITLE Drosophila mitochondrial DNA: a novel gene order
 JOURNAL Nucleic Acids Res. 10 (21), 6619-6637 (1982)
 MEDLINE 83090428
 PUBMED 6294611
 REFERENCE 2 (bases 5269 to 5695)
 AUTHORS Clary,D.O., Wahlthner,J.A. and Wolstenholme,D.R.
 TITLE Transfer RNA genes in Drosophila mitochondrial DNA: related 5' flanking sequences and comparisons to mammalian mitochondrial tRNA genes
 JOURNAL Nucleic Acids Res. 11 (8), 2411-2425 (1983)
 MEDLINE 8320794
 PUBMED 6304652
 REFERENCE 3 (bases 404 to 5272)
 AUTHORS de Bruijn,M.H.
 TITLE Drosophila melanogaster mitochondrial DNA, a novel organization and generic code
 JOURNAL Nature 304 (5923), 234-241 (1983)
 MEDLINE 8324508
 PUBMED 6408489
 REFERENCE 4 (bases 804 to 1778)
 AUTHORS Satta,Y., Ishiwa,H. and Chigusa,S.I.
 TITLE Analysis of nucleotide substitutions of mitochondrial DNAs in Drosophila melanogaster and its sibling species
 JOURNAL Mol. Biol. Evol. 4 (6), 638-650 (1987)
 MEDLINE 88174373
 PUBMED 2832697
 REFERENCE 5 (bases 5268 to 13619)
 AUTHORS Garesse,R.
 TITLE Drosophila melanogaster mitochondrial DNA: gene organization and evolutionary considerations
 JOURNAL Genetics 118 (4), 649-663 (1988)
 MEDLINE 8812147
 PUBMED 3130291
 REFERENCE 6 (bases 441 to 2967)
 AUTHORS Satta,Y. and Takahata,N.
 TITLE Evolution of Drosophila mitochondrial DNA and the history of the melanogaster subgroup.
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 87 (24), 9558-9562 (1990)
 MEDLINE 91088557
 PUBMED 2124697
 REFERENCE 7 (bases 14215 to 14512)
 AUTHORS Ballard,D.W., Olsen,G.J., Faith,D.P., Odgers,W.A., Rowell,D.M. and Atkinson,P.W.
 TITLE Evidence from 12S ribosomal RNA sequences that onychophorans are modified arthropods
 JOURNAL Science 258 (5086), 1345-1348 (1992)
 MEDLINE 9308057
 PUBMED 1455227
 REFERENCE 8 (bases 14917 to 19517)
 AUTHORS Lewis,D.L., Farr,C.L., Farguhar,A.L. and Kaguni,L.S.
 TITLE Sequence, organization, and evolution of the A+T region of Drosophila melanogaster mitochondrial DNA
 JOURNAL Mol. Biol. Evol. 11 (3), 523-538 (1994)
 MEDLINE 94285822
 PUBMED 8015445
 REFERENCE 9 (bases 1 to 408, 13319 to 19517)
 AUTHORS Lewis,D.L., Farr,C.L. and Kaguni,L.S.
 TITLE Drosophila melanogaster mitochondrial DNA: completion of the nucleotide sequence and evolutionary comparisons
 JOURNAL Insect Mol. Biol. 4 (4), 263-278 (1995)
 MEDLINE 96423163

PUBMED 8825764
 REFERENCE 10 (bases 1 to 19517)
 AUTHORS Lewis,D.L., Farr,C.L. and Kaguni,L.S.
 TITLE Direct Submission
 JOURNAL Submitted (03-OCT-1995) Laurie S. Kaguni, Biochemistry Department, Michigan State University, East Lansing, MI 48824-1319, USA
 FEATURES
 source
 1. 19517
 /organism="Drosophila melanogaster"
 /organelle="mitochondrion"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 /note="derived from new and previously submitted sequences; sequence is a composite containing sequences obtained from different Drosophila melanogaster strains"
 1. .65
 /product="tRNA-Ile"
 complement(97. 165)
 /product="tRNA-gln"
 171. .239
 /product="tRNA-Phe"
 240. .1265
 /codon_start=1
 /transl_table=5
 /product="NADH dehydrogenase subunit 2"
 /protein_id="AAC47811.1"
 /db_xref="GI:1166530"
 /translation="MFNNSKILFTIMIGITLITTSNWSGAMNGEINLSFPL
 LSDNNIMSTESLKYELTQVASTVLFSILMKNNNNINFSFMIMSHLL
 LKGAAPFHFPFNMMEGLTMMNALMIMTQKAPMLISYIKYLLISVLSVI
 GAIGAGNOTSLRLKNAFSSINLGMWLSLMSIESIMLLFFYSFSLSVLFEMIF
 KRLHNLQPSFPMNSKIKETPLFMNLSLGGPLPFGPLKVLVIOQLTCONYFMT
 LTMKSTLITTFPLRICYSAFMKNYFENNMMGNNSINVMYMTFFSIRGLFI
 SLRTFM"
 1264. .1329
 /product="tRNA-Trp"
 complement(1322. 1383)
 /product="tRNA-Cys"
 complement(1403. 1468)
 /product="tRNA-Tyr"
 join(1470. 1472,1474. .3009)
 /note="mechanism underlying reading frame shift after first codon uncertain"
 /transl_table=1
 /codon_start=1
 /product="cytochrome c oxidase subunit I"
 /protein_id="AAC47812.2"
 /db_xref="GI:7412848"
 /translation="MSKQMLPSTHNDIGTLPIFGMAAGVNGNSILIRAEHGPG
 ALIDSDIIVIVTAHAFIMFPMVPMIGGGMNVLPLMLGAPDAFRNNMSW
 LIPALSLILNWSSTGISLDRMPLFWSVIVITALILSLPVLAGATLWLTDRINTSF
 VNFITIVINMSSTGISLDRMPLFWSVIVITALILSLPVLAGATLWLTDRINTSF
 FDPAGGGBPIYCHLPWFPGPEVYIILIPGFIISHIIQESKEKTFSLGNIYVM
 LAIILGFIWAHMFPMGVDVPRAYATAMIIAVPTGIKIFSMATLHGLQISYS
 PAIIMAGFVPLFTVGGTLGVIVLVINSSVILIDHTYVVAHFVYLSMGAVFAMAF
 IHVPLPGLTLNKKKSHFIIIMFIVNLTFFPOHFLGAGMRRRSDVDATYNN
 IVSTIGTSLIGLIFFFITLWESLVQRQVITYIQLNSSLIEWQNTPPRHSSELP
 LLIN"
 3012. .3077
 /product="tRNA-Leu"
 3083. .3767
 /note="TAA stop codon is completed by the addition of 3' A residues to the mRNA"
 /transl_except=(pos:3767,aa:TEAA)
 /transl_table=5
 /product="cytochrome c oxidase subunit II"
 /protein_id="AAC47813.1"
 /db_xref="GI:1166532"
 /translation="MSTWANKGLDSASPLMEOLIFPHDHLILVMTVAVGLMFM
 LFFNNYVRLIHGLQIMTWITLPAIILFIALPSRLVILDEINBPSVTLISGH
 QWMSYSDSNNIISFDSYIMPLNMLNDGRLDDVNRVYLPNSQIRILVTAADYI
 HSWIVPALGVKVDGTPRLNQTINFFINPGLFYQCCEBICGANSFMPYIESVPVY

Tue Apr 6 07:57:22 2004

us-10-081-051-52.rge

Page 79

```
Db 15454 TTTTGTAAAAAAATATTTTATATATATTTTCATTTTTTTTTTAAAAAAATTTT 15395
QY 3653 CTTAATCCATCATATGAGTCTCTTAACATATATCCATATATGAAATCAGGCTCT 3712
Db 15394 TTAATAATTTATATTTATATTTTAAAAAATGTAATAAAGTTTAAAAATCTTA 15335
QY 3713 TTTTCATTAAGCCTGATCGGATTAACATCAATGCTAATTAGTGAATATTA 3772
Db 15334 AAAAAAATTTTTTTAGCATTTAAAGAAATTAATCATATTTTTTAAAAATTTT 15275
QY 3773 TCATCACTTAATATCTTTTATATAGATTTCGATATCACTCATCATCAACTAAT 3832
Db 15274 TAAAAAATTTTTTTTTTAAAAAATTTTTTTTTTAAAAAAT - ATTTTTAA 15217
QY 3833 ACTTCAGAAATATACAACTCTTCCTGACATTTCAAAATCCTGTCATTAATTAACCTAC 3892
Db 15216 AAAGTGTAATTAATTAATTTTATTTTAAATTTAAATTTTATTTATATATTTGA 15157
QY 3893 AATTAATGAAATTAATGCAACACA-ATGTAATCTCTTAAGTACACCTCATATATCTA 3951
Db 15156 TAAATTTATTTTTTGTAAAAATTAATTTTTTTTAAAAAATTTTAAAAATTT 15097
QY 3952 CCACACACACTAATGATCAACACCAATATATTAATTCATTTTTTTACTATTAAT 4011
Db 15096 TTGTAAATTTTAAAAATTAATTTTATTAATTTTATATATATATATGATAAAAAT 15037
QY 4012 TAAACATATTAATAAATTAAGTACTAT -ATTACTAATTATCAAGCATAGTTTCT 4069
Db 15036 TTAATTAATAATTAATTTATTAATTAATTAATAATTAATTTTCTTAA 14977
QY 4070 ATTAATCCAAAAAGAAAAATTAATAACCTAATTAATAAATCTATCATCACTTGTAAC 4129
Db 14976 AATATATTTTATTAATAATTAATTAATTAATTAATAATAAATAAATAAAT 14917
QY 4130 CTAATAGTAATAATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4189
Db 14916 TTAAGTTTATTTGGCTTAAAAATTTGTATTAATTAATTAATTAATTAATTTTGT 14857
QY 4190 GTCTATAGTTTTCATTAATCTAAAACTGAGCTTATAGATATTCATATAT 4244
Db 14856 GTGATTTTATTAATTAATTAATAATTAATGAATAATTAATGAGTAATTAAT 14802
```

Search completed: April 5, 2004, 18:26:33
Job time : 18045 secs

Db 776 GCCAAGACCGGTAATAGCGGGTTGTAATATCCATGCGCATATGCCCAATTCATCAAGA 717
QY 3128 AATAACGATACATGATTTGGCTTCTCTATATTCGAATATGAGAGTACTATAGAGC 3187
Db 716 AAGAGGAAACCGCATTTCTGTTCTCAAAACGCGCTGGCGATTTGAGATGACACAGTA 657
QY 3188 AATATATATTTACTTCCCTACCTACCAATATATTAACCAAGTATTTTACGATGAC 3247
Db 656 AATGCCCTTTTTCATGACCAAAAGTTCTGATTCATTAATTAAGTCTTGGGAATAGCGCC 597
QY 3248 ATAGATGATACATATGATTAATGAGATATCATACCCCTTGATTTTATATATAGC 3307
Db 596 ATA-----TTGAGTGCATTAATAAGGGGCTTGCTCGGACTATGGCGATGAAGACA 543
QY 3308 CTAGCACTACTTCTTCTTCCACACCTTGGCGAACCAGTAATGAGTATGACTGATGTA 3367
Db 542 TGTGCAACTAATCTTTTACCACTACCCGATCGCATTTGATTAAGCACTAATGATGAG 483
QY 3368 GTAGCTGCTTTATATCATATCTTCTCAATTAACGATTAACAGGAGTTACCGACTATT 3427
Db 482 CGAGAAAGCGCGCATATATGATTAACCTTGTGATGCGAGGCGCTTCCCAATCATG 423
QY 3428 TCATA--ATCCTCAATGCTGATTTCAATCATATTTTCTTACGTAATCTACAGAC 3484
Db 422 TCAGACACAGATTAAGGTTAGTTCTGTGATTAAGTTGTTTGTCTCGATAGTG 363
QY 3485 TCATATGCTCTCTTCACTCACTTAATCTTCTGTAAGGCTTTCTATATA 3544
Db 362 GATATGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 303
QY 3545 TCATTAAGACCATATGACAGACTTTACAGAGTGGCAATTTACATGCCCACTAATC 3604
Db 302 TCATAAGCCCTTGGCTGATTAACATTTACCGGATCAAGATGAGGCTGTCTCATG 243
QY 3605 ATATATACAGGCAATTAAGGATACCTTCTTAACTTTTCCAGACCTTAATCATCA 3664
Db 242 ATATATACCGGCACTAGGAGTGTGCTCTTTAAGCTTTTAAATAGATTAACATCA 183
QY 3665 ATATCAG 3671
Db 182 ATATCAG 176
RESULT 2
US-09-328-352-3206/c
; Sequence 3206, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GIG99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 3206
; LENGTH: 1557
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-3206
Query Match 3.4%; Score 160; DB 4; Length 1557;
Best Local Similarity 48.3%; Pred. No. 5e-21;
Matches 553; Conservative 0; Mismatches 575; Indels 18; Gaps 3;
QY 2699 TTTCGTATATCTTCTTCTTGAAGATTTCAATTAATAATCAATTCATATACATAGT 2758
Db 1208 TCTTCTGATTAACCTCAGACAGATTAATCAATGACGTTAACCAAGGATATTTTCA 1149
QY 2759 AATGCTTAAAGTACCTGGCATTCATATGACTGATGCTATTAAGCTTCATCACTT 2818
Db 1148 AGCTGACGAACATTAACAGGCTATGTAATGTTGATATAGTCGTTGTTCAAGTGCT 1089

QY 2819 AATACATGATACACAACCTATTTTTCACAGATGCTATGATTAATAATACCTACACAT 2878
Db 1088 AATATCTTGGATTGAGCTTACGCTTCTTTCACCGCGCGGCAAGAAATGTTGAGCGAC 1029
QY 2879 TCCGATATATCTGTACATATTTCTACTAAAGACGATCTTAATTTGAAGACATTAAT 2938
Db 1028 ATCGGATGCTTCACTGAGATGCGCAAGCTCGGATATGATTCGAATTAACATTTAGA 969
QY 2939 CTATTAATTAATCTTCAACAAACCTACAGCTTTTACTTCACTTTCATCTTTGGA 2998
Db 968 CGGTGTATTAAGCTTTCACGGAACACACTTATTCACAGTTTTCATGATCTTGATGCG 909
QY 2999 GAAGACACATTAATTTCTACATCTATACAGATCTTATATTTTCCCTATATAT 3058
Db 908 GTTCCGCGACATTTCTGACATCTTATTAAGGGATATGTCACCAACTCGGTAATAAC 849
QY 3059 TTTCCTCTGTATATATCTGATTAATCTTAATTTGCTATGCTATGCTAAATCTGTACT 3118
Db 848 TCACCATCTGCGACACAGTAAACATCGAGTTGTGTTCAAGGCGATGACCAAT 789
QY 3119 TCATCTATAATTAACGATCAATGATTTGCTGCTCTATATTTCAATATGAGAGGATCT 3178
Db 788 TCATCAAGATATGTTTCCGCAATTTGCTTGTGTTCAAGGCGCTTATGTTGGGTTT 729
QY 3179 CTATGACATATATTAATTTACTTTCTCACTACCAATATATTAACCAATATTA 3238
Db 728 GCACCGGTGA-----AGACCTTTTTCATGCAAAATATTCGTTTCGATTAATCT 675
QY 3239 GCTGTAGCATAGATGATGATATATTAATGAGATATCAATCCCTGTGATTTT 3298
Db 674 TTAGATATGCTCCATGTTTAAAGCAATTAAGCTTACCGCGTGTGTATGTT 615
QY 3299 TGTATTAAGCTAGCACTATCTTCTTCAACACCTGCGAAACGATTAATGATATGCA 3358
Db 614 TGTAAAGCATGTCACACAGTTCTTACCTGTGCTGATTCACATTAATTAACATGTA 555
QY 3359 CTGATGATGATGCTCTTATTAATCACTTCTCAATTAATGATTAACAGGAGTGA 3418
Db 554 ATATGCGACTGATTAAGCAACCAATCGCTCGGAAACCTTTGCTATACAGAGATG 495
QY 3419 CGCATATTTTCATATCTCA-----ATGCTGATTTCACTCATCTATTTCTAGCT 3472
Db 494 CCAATATTTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 435
QY 3473 AATCTACAGACTCTATAGCTCTTTTACACATTAATCTTCTCTGTAAGAGC 3532
Db 434 TATATGATATGAATTAATGCTGATGATTAAGCTTAAGCTTCAATATGATGAT 375
QY 3533 TTTTCTATATATCAATAGACCAATATGACAGCTTTTACAGAGTGGCAATATTACA 3592
Db 374 TTGTGTAATATCAATGACAGCTGATGATGATGATGATGATGATGATGATGATGAT 315
QY 3593 TGGCCATATCAATTAAGAGCAATTAAGATTAATCTTTCTTAACTTTTCCAGTACA 3652
Db 314 TGGCTGTCTATTAATGAATGAAGCAAGTGGGTGAGATTTTCACTTTAGATTAAGAA 255
QY 3653 CTATATCATCAATATCAATGATCTCTTAAACATATATCAATTAATCAATCAATGAGCT 3712
Db 254 GTTAACCATCAATA-----CTGGCATATGCAATGATTAAGATGATGATGATGATGAT 201
QY 3713 TTTTCTATTAAGCACTTTGATGCGGATTAACATCAATGCTATTAATTAAGATATTA 3772
Db 200 TCATGATGTAACGTTCAAGTGTGTTGATGCTTTCAAAATTTGAGCTAAATCT 141
QY 3773 TCATCACTATATATCTTTTATTAATTTCTGATATCAATCAATCAATCAATCAATCAAT 3832
Db 140 TCTTCTTAAAGCTTTTCAAGTACCATGATGATGATGATGATGATGATGATGATGAT 81
QY 3833 ACTTCA 3838
Db 80 ATTTTA 75

RESULT 3

US-09-107-532A-350
Sequence 350, Application US/09107532A
Patent No. 6583275

GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 7310

CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354

COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>

CURRENT APPLICATION DATA:
SOFTWARE: ASCII

APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998

APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:
NAME: Ariehello, Pamela Deneke

REGISTRATION NUMBER: 40,489

REFERENCE/DOCKET NUMBER: GTC-012

TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007

TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 350:

SEQUENCE CHARACTERISTICS:

LENGTH: 1071 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: circular

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Enterococcus faecium

FEATURE:

NAME/KEY: misc_feature

LOCATION: (B) LOCATION 1...1071

SEQUENCE DESCRIPTION: SEQ ID NO: 350:

US-09-107-532A-350

Query Match 3.3%, Score 155, DB 4, Length 1071,
Best Local Similarity 49.0%, Pred. No. 3,98-20,
Matches 504, Conservative 0, Mismatches 515, Indels 10, Gaps 3,

Db 584 AATAATATATATATTTGACTGTTCTTGATATATATTTAGTAAAAATGAAATCAGT 643
Db 10 ACTATATAGGAATATATATATTTATTTAGAGAAATATCCATGAAACAAAAATGCA 69
Qy 644 ATTTAGGTGACGATCATTTTGGCAACGCAATAGCAATGACCTGTAGACATGATATA 703
Db 70 GTTCTGTGCTGATCTTGGGGACATGCTTTAGCAAAACATTTGGAGAAAACGGGCAT 129
Qy 704 TCAGTTAATCTTATGGGACGTGATCATGAAATATTTACATATATAACCTTACCGAAA 763
Db 130 GATGCGCATTTTGGGAAATGTACTGACAGATGATGAAATCAATAGTATCATACA 189
Qy 764 AATTAAATATTTTACCAATATCATTTACCAACAAATATATGCAACCAAGAAATATA 823
Db 130 AATGACATTTTATTAACCTGATTTGACGATCCCAAGATCGATTATGATATTAAGAAATTG 249

Qy 824 GACGAGATTTATCTGACACAAATAC--ATGATATATCTTAATCTTCCATACAAACAT 881
Db 250 GCAGAGCAGTGTATACGACAGATGCTATTTATTTGTTGTTCCAAACAAAGCGATTCGT 309
Qy 882 TAGCAGCATATGTATACAAATATACACAAAGACGATATGTAAAAATCTCAATAT 941
Db 310 TCAGTCGACGAACTTATTTGTTAAATGAAATCTAAGCCATATATACAGCAAGT 369
Qy 942 TAATTTAGTAAAGATATCGAAATTAATCATCACTCAAAATTTCCAGTAAATAGCAAG 1001
Db 370 AAGGATTAAGAACAGATACGATTAAGAAATCCAGAAATTTGGCTGAAGAAATTCGA 429
Qy 1002 AATTTTCAATATATATCAATTTTATATCTCTGCTGGCCAGTTTGGCTAAAGAAATTG 1061
Db 430 GAA--GAAAAAGACAGCAATTTGTTATTTATCTGCTTACGCAATGCTGAGGAAGTGG 487
Qy 1062 CAGAACATCTTCTTGTAGTATAGTACTGCTGCTGTGATATTAAGAACTTGTGATCAT 1121
Db 488 CAGTACATGATATACAAACATTTACAGCTGCAAGATTTGATCAAAAGACAGCGCTATG 547
Qy 1122 TGATGAAACATATATATATATATGTTCTTAAATATATATACATCAAGATTTATAGTG 1181
Db 548 TACAAAACTGTTCAATGATATATATTTACAGATTTTACAAATATGATGTGATGGCG 607
Qy 1182 TACAGATTGAGCTGATTAAGAAACATATATTTGCAATTTGCAATTTGATTAATCGCTGGA 1241
Db 608 TGAATCTGCTGACACTGAAAAATATATATGATTTGCTGCGGCGCTATTCATGAT 667
Qy 1242 AATATTTAGTATATATCTGTTGCTGCTGTTATATCTTAAGCATGATGAATTAATA 1301
Db 668 TAGAATCGAGATGATCAAAAGACGATCATGACCTGCGGATTTGCAAAATATCACCC 727
Qy 1302 CACTATATATGCAAAAAATATCATTCATATATATATATATATATATATATATATATAT 1361
Db 728 GTCTTGTTGCTGCGATGGT-----GCCAATCTCTGACATTTATGCTTTGATGATG 781
Qy 1362 TTGAGATCTATATATATATATATATATATATATATATATATATATATATATATAT 1421
Db 782 TGGGGACTGATCTGTTACTTCTTCAAGTGTTCATCTCGCAATTTGCGGCTCGGTAAT 841
Qy 1422 AATATGAAAAAGTAAATATATATATATATATATATATATATATATATATATATATAT 1481
Db 842 TGTTAGGACAGGACATATAGCTGGAAGATTTCTAGAAAAATATGGAATGCTGTTGAG 901
Qy 1482 GAACAGTACTGTAAACCACTGATATCATTTAGCAAAAAATTAATATAGTATACCA 1541
Db 902 GTGTGGCAACGACAAAGCAGAGTAGCTACACAGCTAGGTGTTGAGATGCCAA 961
Qy 1542 TTTCATATCTTATTTCAATTTATTTATCTGAGAAATATATCTAGTATTAACCATATCA 1601
Db 962 TTACTCAACATCTATATATGTTTGTATATGAGAAATATCAAAAAACAGCAAAAG 1021
Qy 1602 ACATATAT 1610
Db 1022 AATCATGT 1030

RESULT 4

US-08-956-171E-513/C
Sequence 513, Application US/08956171E
Patent No. 6593114

GENERAL INFORMATION:

APPLICANT: Charles Kunsch
Gill H. Choi
Patrick S. Dillon

Craig A. Rosen
Steven C. Barash

Michael R. Fannon

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences

NUMBER OF SEQUENCES: 5256

CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue

CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997

ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439

INFORMATION FOR SEQ ID NO: 513:
SEQUENCE CHARACTERISTICS:
LENGTH: 918 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 513:
US-08-956-171E-513

Query Match 3.2%; Score 154.2; DB 4; Length 918;
Best Local Similarity 53.6%; Pred. No. 5.4e-20;
Matches 346; Conservative 0; Mismatches 293; Indels 6; Gaps 1;

DB 1022 ATTTTATACCTCTGCTGCTCAAGTTTGGCTAAAGAAATGCGAAGATCTTCTGTGTGT 1081
DB 732 ATTGGCGCTGTGTCAGGCGCACTCATGCGAAGAGTTGTATCTCAGCAACCACTACA 673
DB 1082 ATAGTACTGCTGCTGATATAAAGAACTTGTGAATCATTTGATAGAAACAATAAGTAT 1141
DB 672 GTTGGCTGCTTCATCAAAAGATTAAGTATGATTAATTAACGCAAGATTTATTTAGTAA 613
DB 1142 GATGCTCTAAATAATATATACATCAAGATATATAGTGTACAGTTCAGATCTGCATTA 1201
DB 612 GATTATTTGGGTGTACACGAATGATGACTGATTTGGTGTGAACTTGTGTGCTGATTTG 553
DB 1202 AAGACATATATTCGAATTGCTGATGAAATATGCTGGAAGAAATTTAGTATATATGCT 1261
DB 552 AAAAATATCTCCAGAGCAAGTGTGATGTGCTGGAATTTGGTACCGTATATATGCA 493
DB 1262 GTTGTCTACTTATTAACATAAGCAATGAAATTAACATACTATATACCAAAAAAT 1321
DB 492 AAAGCTGCAATTAATGACTGCTGCTTACCGGAATTAATGTA-----TTAGGTGAAG 439
DB 1322 CATTCATATAGATCTTCAATTAATTTGGTCTCATCTGCTTGAAGATCTAATATTAACA 1381
DB 438 TTAGGTCCGATCTTATGACATTTCTAGGTTAGGTGATGATGATGATGATGATGATGAT 379
DB 1382 TGTACACAGAAATTCACCGCAATATGCTTTGGCTAGTAAGTAATAGAAAGAGTAAAT 1441
DB 378 TGACATCAACACATCTCTGGAATTTCACTTAGGATATATTAACCTTGGACAAAGTATCA 319
DB 1442 ATAAATCATTAATATAGATCAACCTTAAGCTTGTGGAAGAAACCGATCTGTAACCA 1501
DB 318 ATGATCAAGCATTTATCTGAATATGATATGATGATGATGATGATGATGATGATGATGAT 259
DB 1502 CTGATATCATTTAGCAAAAAAATTAATGATGAACTCAATTTGCAATATCTATTTACAT 1561
DB 258 GTTATCATTTAGCTAAAGAAAAAATGATGATGATGATGATGATGATGATGATGATGATGAT 199

DB 1562 TTATACATGAGATATATATCACTAGATTAAGCAATATATATATCTAGTCTATC 1621
DB 198 GTATTAATTTGAATATATCTCAGTAAAGAAATGCGTAAAGATTTATGAGCCCGCATAA 139
DB 1622 ATCTTGTATTAATTTCAACCAATGATTAATTAACGCTCAAAAT 1666
DB 138 AATCTGAATTAATGATGATTAATTTACATATACGCTAATATAT 94

RESULT 5
US-09-107-532A-1399
Sequence 1399, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354

COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneka
REGISTRATION NUMBER: 40,483
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 893-5007
TELEFAX: (781) 893-8277

INFORMATION FOR SEQ ID NO: 1399:
SEQUENCE CHARACTERISTICS:
LENGTH: 1062 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHEICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium

FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...1062
SEQUENCE DESCRIPTION: SEQ ID NO: 1399:
US-09-107-532A-1399

Query Match 2.9%; Score 139.8; DB 4; Length 1062;
Best Local Similarity 49.6%; Pred. No. 2.5e-17;
Matches 484; Conservative 0; Mismatches 467; Indels 24; Gaps 4;

DB 645 TTTTAGGTGAGATATCTTTGGCAAGCAATGCAATTTGCACTGTGCAAGCATGATATAT 704
DB 56 TTTTAGGACCAAGGCTCTTGGGGTACTCTTTGGCAAGTCTGCGGAAAGCGACATG 115
DB 705 CAGTTAATCTATGGGACGATCATGAATATTTACATATTAACACTTACCGAAAA 764

Tue Apr 6 07:57:22 2004

us-10-081-051-52.rn1

Page 5

116 AGTTCGATTTGGGGACATCGTCTCAAGTAGAAGAAATCAACCCACACAGCA 175
176 ATTTAAATATTTACCCACATATCATACACAGACAAATATATGCAACCGAATATAG 824
176 ATCAAGATTTATTCAGATCATATATATCCGATTCATCCAGATTTACCAATATAG 235
825 ACAGATATTTATCTGACAAACATATATATTTATCTTACATTTCCACACATATAC 884
236 AAGTAGAGATTAAGAGCGAGATGCG--GAGTCTATTCGATTTCAACGAAAGCAATAC 292
885 GACCATATATGACAAATATCAACAAACAGCATATGTGTAATATCTCAATATTA 944
293 GTTCTGTTGCAAAAGACTAGTCTTAA-----TTGATATCAAGCCAGTATCATCC 346
945 TTTGATAGTAAAGTATCGAATATATATATATATTTCCAGTAAATAGCAGAGA-- 1002
347 ATGCAAGTAAAGATTAAGAACAGGGAACATATAAGATCTTACATCTGTTAGAG 406
1003 -----AATTTACATATATTCATTTTATATCTCTGTTGCAAGTTTGTAAAG 1035
407 AATCCAGCAGAAAAAAGAGAGTGTGTTGTTCTTTCTGACCAAGCCATGAGAG 466
1056 AATAGCAGACATCTCTCTGATATATATATATATATATATATATATATATATAT 1115
467 AAGTAGCAAAATATGACATCATATGATCAGATCAGCTGCAAGTACAGATGCAAGAGCA 526
1116 AATCATTTAGTAAACATATATATATATATATATATATATATATATATATATAT 1175
527 AATATGATACAAAGCTGTTATGATATATATATATATATATATATATATATATAT 586
1176 TAGGTGATACAGATTTGAGAGTGCATTAAGAACATATATGCAATGCAATGGAATAT 1235
587 TCGGTGATAGATGAGGAGCTTTGAAATATATATATATATATATATATATATATAT 646
1236 CTGGAATAATTTAGTAAATATATATATATATATATATATATATATATATATAT 1295
647 ACGGCTTAGGTTTGGCGATATATGCAAGGACGATATATGCGCGGATTAGCTGAGA 706
1296 TTTAAACATATATATATATATATATATATATATATATATATATATATATATAT 1355
707 TCGCCGTTAGTGTAGATATGAGTGC-----ACCGTGAACCTTTATCGATTA 760
1356 CATGCTTGGAGATCTATATATATATATATATATATATATATATATATATATAT 1415
761 GGGAAATAGGCGATCTAGTCTTATGCAAGTGTCTTATGGAATGCGCGCGAG 820
1416 GACTAGAAATAGAAAAAGGTAGAAATATATATATATATATATATATATATATAT 1475
821 GGAATCTGTTAGAACAAAGGAGCGCCCTAGAGAGGTCTAGCCAAATGGAATGCGC 880
1476 TTGAAGAAACCGATCTGTAACACCATGATATCTTAGCAAAAAAATTATATAGAAC 1535
881 TTGAAGGTGTGATATCTACCAAGAGGATGGAATGAGCCCAATATATATATATAT 940
1536 TACCAATTTGATATCTATTTTCAATTTATATGAGATATATATATATATATATAT 1595
941 TCCCATCATCTCAAAACATCTATGAGAGTGTATGAACAAAAAAGAAATCAAAACAG 1000
1596 TATCAACATATAT 1610
1001 CAAAGAAATCATGT 1015

RESULT 6
US-09-557-884-1
Sequence 1, Application US/09557884
Patent No. 6506581
GENERAL INFORMATION:
APPLICANT: Fleischmann et al.
TITLE OF INVENTION: The Nucleotide sequence of
the Haemophilus influenzae Rd Genome, Fragments
Thereof, and Uses Thereof
NUMBER OF SEQUENCES: 1

CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS V6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/557,884
FILING DATE: 25-Apr-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/476,102
FILING DATE: JUN-5-1995
ATTORNEY/AGENT INFORMATION:
NAME: Michelle S. Marks
REGISTRATION NUMBER: 41,971
REFERENCE/DOCKET NUMBER: PB186P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1830121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1
Query Match 2.9%; Score 136.8; DB 4; Length 1830121;
Best Local Similarity 50.4%; Pred. No. 3.1e-16;
Matches 476; Conservative 0; Mismatches 447; Indels 21; Gaps 5;
620 AATAGTAAATAATGAAATACATATTTAGTGTGACGATCATTTGGCACAGCAATAGCA 679
635171 ATAACTTGGCAACACCAATCACTGTGCTAGGTCGGGATCTTATGAACTGCGTGGCA 635230
680 ATGCACTGACACACATGATATATATATATATATATATATATATATATATATATAT 739
635231 ATTACCTTTCTGCGCATGTTCTCCACTACTTATGGGGGACAAATCCCGCCATATC 635290
740 ACAATATATTAACCTTACCGAAAAATTTAAATATTTTACCACATATATATATATAT 799
635291 GCAAAATAGGAGACAAAGCAAAATATATATATATATATATATATATATATATAT 859
800 AACATATATGCAACCGACATATATAGACAGATATATCTGACACATATATATATAT 919
635351 GATCTTCATTTAGAAAGTATCTTGCAAGCAAT--GGAATATCTCAAGATATTTTA 940
860 TTAATATTTCTTACACAAATATGCGACCATATGTTACCAATATACAAACAAACAGAT 999
635408 ATCGGTGTGCAAGCATGTTCTGTTGAATTTCTATATAAT-----CAAACCGCAC 1058
920 ATGTGTAATAATCTGCATATATATATATATATATATATATATATATATATATATAT 1117
635462 TTTAAAGCCCATACCGATTTATTTTGGGCAAAAGTTTGGAGCTTAACAGGCGCT 1176
977 AATTTCCAGTAAATAGCAAGAAATTTTACATATATATATATATATATATATATATAT 1235
635522 TTAATCAAAACGTTAGTAAAGCAACATCGGACGATATATATATATATATATATAT 1294
1037 GGTCCAGATTTGCTTAAAGAAATTTGCAAGCAATCTCTGTTAGTATATATATATATAT 1353
635582 GGGCCCACTTTTGGCAAAAGAAATTTGCAAGGCTTACGATGCTTTAGCTGCTGCC 1412
1097 GATATATTAAGAACTTGT--GTGAATCATTTAGTAAAGAAATATATATATATATATAT 1471

Db 635642 AATAAGAGCAATTCGACGGGAATTCAGTCTGATTCATGTAAGATTCCGA 635701
Qy 1154 AATAATACATCAAGATATTTAGGTGACAGATTGGAGCATTAAGAATTAAT 1213
Db 635702 GTTTATATAATTCGATATGACGTGGCTTCACTGGTGAGCAATTAATAATGATC 635761
Qy 1214 GCATTGCAATGGAATTAATCGTCGAAAAAATTTAGTAATATGCTGTTGCTGTT 1273
Db 635762 GCATTGTCGGGTATTTAGACGGTATGAGATTGGCGCAATCTCGCACAGACTG 635821
Qy 1274 ATAACTAAGAGCATGAATGAATTAACAATATATAGCAAAAAATCATTAATAGT 1333
Db 635822 ATTACTGAGGTATTTGAGAAATTAACCCGCTAGTATCTC-----GCTCGGTGCAAT 635875
Qy 1334 CTTCATCATTAATTTGTCATCATGCTTGGAGCTAATTAATCATGTCACAGAA 1393
Db 635876 ACCAATACATTTATGGGAATGCTGATTTGGCGATTTAGTCTTACCCGATAT 635935
Qy 1394 CATTACGCAATATGCTTTTGGACTAGAAATAGAAAAAGTAGAAATATAATACATTA 1453
Db 635936 CAATCGGCAATGAGCTTTTGGATTAATGCTAGTAAGATTAATGATGCCAAATGGCT 635995
Qy 1454 ATAGATACAACTTAAGCTTTGTAAGGACAGACTGTAATAACCATGATATCATTA 1513
Db 635996 ATGGAATAATATCGGGCAAGTCGTAGAAAGGTTTATATACAAAGAACCTTATTTGCTT 636055
Qy 1514 GCAAAAAAAGCTTAATGTAAGTAACCAATTTGCATATCTATTTA 1557
Db 636056 GCAAAAGACAGAGTGTGAAATGCGATACAGAACAAATTTA 636099

RESULT 7
US-09-643-990A-1
Sequence 1, Application US/09643990A
Patent No. 6528289

GENERAL INFORMATION:

APPLICANT: Robert D. Fleischmann

Mark D. Adams

Owen White

Hamilton O. Smith

J. Craig Venter

TITLE OF INVENTION: The Nucleotide sequence of the Haemophilus influenzae Rd Genome, Fragments Thereof, and Uses Thereof

NUMBER OF SEQUENCES: 1

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville,

STATE: MD

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: 3 1/2 inch diskette

COMPUTER: Dell Pentium

OPERATING SYSTEM: MS DOS V6.22

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/643,990A

FILING DATE: 23-Aug-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/487,429

FILING DATE: 1995-06-07

APPLICATION NUMBER: 08/426,787

FILING DATE: 1995-04-21

ATTORNEY/AGENT INFORMATION:

NAME: Kenley K. Hoover

REGISTRATION NUMBER: 40,302

REFERENCE/DOCKET NUMBER: PB186P1C1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 301-610-5790

TELEFAX: 310-309-8439

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1830121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1

Query Match 2.9%; Score 136.8; DB 4; Length 1830121;
Best Local Similarity 50.4%; Pred. No. 3.1e-16;
Matches 476; Conservative 0; Mismatches 447; Indels 21; Gaps 5;

Qy 620 ATTAGTAAATAAGAAATCAGTATTTAGTGCAGATCATTTGGCAGCAATAGCA 679
Db 635171 ATACTTCGCAACACCAATCATCTGCTAGTGGGAGCTTATAGAACTGGGCTGGCA 635230
Qy 680 ATTGACTGTCAGCAATGCTATATCACTTAATGAGGAGCTGATCATAGAAATTT 739
Db 635231 ATTACTTTTCTCCGAATGCTTCCCACTCATTAATGAGGAGCAATCCGCCCATATC 635290
Qy 740 AACATATTAACCTTACCCGAAAAATTTAAATATTACCAATATCATATACAGAC 799
Db 635291 GCACAATTCAGACAGAACAGCAAAATTAATGCTTTTGGCGGATGTCATCTTCTGAG 635350
Qy 800 AACATATTAACCAACCAATATAGAGAAATATATCTGACCAACAAATACATGTTATC 859
Db 635351 GATCTTCAATTAAGAAATATCTTGACAGCAAT---GGAATATCTCAAGATATTTTA 635407
Qy 860 TTACTATTTCTACACAAATATGACCAATATGACAAATACACAAACCAAGCAT 929
Db 635408 ATCGTGCGCAACCCATGCTTTCGGTGAATTTCTTAATAAT-----CAACCGCAC 635461
Qy 920 ATGTGTAAATAATCTCCAAATATTAATTTAGTAAGATATGAA---ATTACATCATC 976
Db 635462 TTAAGGCCATCATCCGATTAATTTGGCAACAAAGTTTGGAACTACACAGGGCGT 635521
Qy 977 AAATTTCCAGTAAATAGCAAGAAATTTTACAAATATCAATTTTATATCTCT 1036
Db 635522 TTACTACAAACCGAGTGTGAAGAACTCGGAACGAATATCCTAGCTGATCTTCT 635581
Qy 1037 GGTCCAGTTTGTGTAAGAAATTCAGAAATCTTCTTGTATAGTATGCTGCT 1096
Db 635582 GGCCCACTTTTGCAAAAGAAATTTGCGCAAGTCTACATCAGATATTAAGCTGTGCC 635641
Qy 1097 GATTAATAAGACTTG---GTGAATCTTGTATAGAAACAAATAGTATGATGTTCTAAA 1153
Db 635642 AATAAGCAATTCGACGGGAATTTCACTGCTGTTTATGTTATGTAAGAGTTTCCGA 635701
Qy 1154 AATAATACATCAAGATATTTAGGTGACAGATTGGAGCATTAAGAATTAAT 1213
Db 635702 GTTTATATAATTCGATATGACGTGGCTTCACTGGTGAGCAATTAATAATGATC 635761
Qy 1214 GCATTGCAATGGAATTAATCGTCGAAAAAATTTAGTAATATGCTGTTGCTGTT 1273
Db 635762 GCATTGTCGGGTATTTAGACGGTATGAGATTGGCGCAATCTCGCACAGACTG 635821
Qy 1274 ATAACTAAGAGCATGAATGAATTAACAATATATAGCAAAAAATCATTAATAGT 1333
Db 635822 ATTACTGAGGTATTTGAGAAATTAACCCGCTAGTATCTC-----GCTCGGTGCAAT 635875
Qy 1334 CTTCATCATTAATTTGTCATCATGCTTGGAGCTAATTAATCATGTCACAGAA 1393
Db 635876 ACCAATACATTTATGGGAATGCTGATTTGGCGATTTAGTCTTACCCGATAT 635935
Qy 1394 CATTACGCAATATGCTTTTGGACTAGAAATAGAAAAAGTAGAAATATAATACATTA 1453
Db 635936 CAATCGGCAATGAGCTTTTGGATTAATGCTAGTAAGATTAATGATGCCAAATGGCT 635995
Qy 1454 ATAGATACAACTTAAGCTTTGTAAGGACAGACTGTAATAACCATGATATCATTA 1513
Db 635996 ATGGAATAATATCGGGCAAGTCGTAGAAAGGTTTATATATCAAAAGAACCTTATTTGCTT 636055

QY 1514 GCAAAAACTTATAGTACCACTACCAATTGCAATCTATCTATTTA 1557
DB 636056 GCACAAAGACAGAGTGTGGAAATGCCATTTACAGACAAATTTA 636059

RESULT 8
US-09-134-000C-1318
; Sequence 1318, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1318
; LENGTH: 999
; TYPE: DNA
; ORGANISM: Enterococcus faecalis
US-09-134-000C-1318

Query Match 2.8%; Score 134.6; DB 4; Length 999;
Best Local Similarity 48.5%; Pred. No. 2.3e-16;
Matches 465; Conservative 0; Mismatches 484; Indels 10; Gaps 3;

QY 654 CAGATCATTTGGGACGACGATGACATTTGACGTCGTGACGACGATGATTCAGTTACT 713
DB 5 CAGGCTCTGGGGGACGCTTTAGTTCAAGTATTAGCCGAAATGGACGAGGCTGTGA 64
QY 714 TATGGGACGTCATCAGAAATTTACATATTAACCTTACCGAAAAATTTAAAT 773
DB 65 TTGGGGAAATTAACGAGCAATGATGAATTAATTAACCAACACCAACAAACATT 124
QY 774 ATTACCCATATCATCTACGACACATATATGCAACGACATATGACGAATAT 833
DB 125 ATTACCAAGATTAATCTACTCAATTCAGGACACGATGATTCAGCACTGCTT 184
QY 834 TATCGACACATATCATCTATCTTA--ACTATCTACACAAATTAACGACAT 891
DB 185 TAGTAGTGTGATGCGGTACTTTGTGTGACCACTAAAGCATTGCTGATGACAC 244
QY 892 ATGTACCAATATACACAAACGACATATGTGTAATAATCTCCATATTAATTGTAG 951
DB 245 AAGAGTGGCCCACTTTAAAAACAAACCAATTAATTCATGACAGTAAAGGCTCG 304
QY 952 TAAAGTATGAAATTAATCATCAATTTCCAGTGAATAGCAGAAATTTTACA 1011
DB 305 AACAGGTAACATTAAGCGATTTACAGAGTATAGCGAAGAG--ATTCCAGCAAAA 362
QY 1012 ATATATTCATTTTAACTCTGTCGCAAGTTTGTGTAAGAAATTTGACAGACATCT 1071
DB 363 ACGGCAAGGATGTGTTTATCTGTCACAGGCAAGCAAGAGTCCGCTGTATGA 422
QY 1072 TCCCTAGTATATCTGCTGTGTATTAAGAACTGTTGTAATCATGATAGAAAC 1131
DB 423 CATTAACCATTTACAGACAGACGAAATTTAGCAGACGCTGTATGTGCAAGATT 482
QY 1132 AATAGTATGATGTCTAAATATATATATATATATATATATATATATATATATAT 1191
DB 483 GTTATATATGATGATCTTAAATTTATATATATATATATATATATATATATATAT 542
QY 1192 AGCTGCAATTAAGAACTATTTGCAATTTGCAATTTGCAATTTGCAATTTGCAATTT 1251
DB 543 GGGGCTTAAATATTTATTTGTTGGGGCGGAGCAATTCATGCTTAAAGCTTTGG 602
QY 1252 TATATATGCTGTCTACTGTTATATATATATATATATATATATATATATATATAT 1311
DB 603 TGAATATGCAAGGCAATTAATGATGCTGTGTTAGCAGAAATATATATATATATAT 662

QY 1312 AGCAAAATATCATATATATATATATATATATATATATATATATATATATATATAT 1371
DB 663 TGGGATGGG-----AGCAAAATATATATATATATATATATATATATATATATAT 716
QY 1372 AATATTAATGATGACAGACAACTTACGCAATATGCTTTTGAATATAGAA 1431
DB 717 AATGTTATCTGTAAGATGTCATTTTCAAAATGCGGTGCGGTAATTTATAGAAA 776
QY 1432 AGTGAATATATATATATATATATATATATATATATATATATATATATATATAT 1491
DB 777 AGCCATTAAGTATGATGAGTCTTGAAGAAACATGGGCAATGATGAGAGGTTTCAAC 836
QY 1492 TGTAAACCATGATATATATATATATATATATATATATATATATATATATATAT 1551
DB 837 AACTAAAGCAAGCTATGATGATGACCAACCACTGAATGCAATGCAATGCAAGAAC 896
QY 1552 TATTTACATTTTATATATATATATATATATATATATATATATATATATATAT 1610
DB 897 AATCTACATGCTTGTCAATGACGAATGTACACACAGCAAGAAATCATGT 955

RESULT 9
US-09-134-001C-300
; Sequence 300, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 300
; LENGTH: 1059
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-300

Query Match 2.5%; Score 121; DB 4; Length 1059;
Best Local Similarity 47.6%; Pred. No. 7.4e-14;
Matches 460; Conservative 0; Mismatches 495; Indels 12; Gaps 3;

QY 628 AAAATGAAATCAATTTTATGATGACGATCAATTTGCAAGCAATGCAATGCACT 687
DB 60 AATGAGAAATTAACAGTTTGTATGATGATGATTTGTATGATGATGATGATGAT 119
QY 688 GTGACCAATGATATATATATATATATATATATATATATATATATATATATATAT 747
DB 120 AGCTCAAAATGCTATGATGATTTATGATGATGATGATGATGATGATGATGAT 179
QY 748 AAACACTTACGCAAAATTTAAATATTTTACCAATATATATATATATATATATATAT 807
DB 180 TAAACACATCATATATGACAAATATATCTTAAGATGCTAAATATATATATATATAT 239
QY 808 TGCAACCAAGCATATAGCGAAGTATTTCTGACACACATATCATGATATATATATAT 867
DB 240 AGCAACTGTATATATATATATATATATATATATATATATATATATATATATATAT 927
QY 868 TCCCTACACAAATTAACGACATATGATGACAAATATACAAACAGCAATATGATGAT 927
DB 297 AACTTCAAAAGCAATTAAGAGATATCAAAAGATATGATGATATATATATATATATAT 356
QY 928 AATATCTCAATATATATATATATATATATATATATATATATATATATATATATAT 987
DB 357 GACTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 416
QY 988 TGAATATGCAAGAAATTTTACAAATATAT--CAATTTTATATCTCTGCTGACAG 1044

Db 2285 ATGCAGAT-----ANTTAAATTCATCTTAACTCCACACAAATTAATTAATA 2232
 QY 1241 AAAATTTAGGTAATATGCTGTGCTGCTATTAAGGATGATGAATTAATA 1300
 Db 2231 CATATTAATTTAACTCTTACTCTCAAAAACCTAATTTAAATTAACATATTAATA 2172
 QY 1301 ACATATATATGCAAAAATTCATCAATAGATCTTCAATTAATTTGGTCAATGT 1360
 Db 2171 AACACCTAATAATTAATTAATAAAAAAAATCTTAAACCTTAAATTAATA 2112
 QY 1361 CTGGAGATCTAATTAATTAATGATGACAGACATTCAGCAATATGGCTTTGACATA 1420
 Db 2111 TATATAAAATTAATTCCTTAATACATTAATTAATTAATTAATTAATTAATA 2052
 QY 1421 GAAATAGAAAAGTAGAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1480
 Db 2051 AAACAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1594
 QY 1481 GGAACAGATCTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1540
 Db 1993 AAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1534
 QY 1541 ATTGCAATATCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1600
 Db 1933 ATTA-----ACACACACACAGAAAACCAATTAATTAATTAATTAATTA 1878
 QY 1601 AACATATATCTTATGCTATCATCTTGTGATTAATTTCAACACATGATTAATTAACGCTC 1660
 Db 1877 ACCCTATTAACCAATCAAAATCTCTTCCAAATGACACTAAA--AACTTTAAAAAAA 1820
 QY 1661 AAAATATAGATCAACGCTTACACATATTAATTAATTAATTAATTAATTAATTA 1720
 Db 1819 AAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1760
 QY 1721 CATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1780
 Db 1759 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1700
 QY 1781 GCTATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1836
 Db 1699 AATCATTAAGAAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1640
 QY 1837 TTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1896
 Db 1639 TCACCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1580
 QY 1897 TCATCATCATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1956
 Db 1579 TTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1520
 QY 1957 ATACATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2016
 Db 1519 CTATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1460
 QY 2017 CATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2076
 Db 1459 CTAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1400
 QY 2077 TATGCCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2136
 Db 1399 AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1340
 QY 2137 TGCAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2196
 Db 1339 AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1280
 QY 2197 CCACCTACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2256
 Db 1279 CCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1220
 QY 2257 CATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2316
 Db 1219 CTCTACCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1160

QY 2317 CTTCACACAAACTCAAGTTCTATTATTAACAATTCAGAAATTAATTAATTAATTA 2376
 Db 1159 ATCAAAATCCCTCCCTCCCAATCTACAAACCCCAAAAAAATTAATTAATTA 1100
 QY 2377 AACACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2436
 Db 1099 ACCCCAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1042
 QY 2437 TTCATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2496
 Db 1041 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 982
 QY 2497 TTCA 2501
 Db 981 TCTAA 977
 RESULT 11
 US-10-204-708-6/c
 ; Sequence 6, Application US/10204708
 ; Patent No. 6677731
 ; GENERAL INFORMATION:
 ; APPLICANT: OLEK, Alexander
 ; APPLICANT: PIEPENROCK, Christian
 ; APPLICANT: BERLIN, Kurt
 ; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
 ; TITLE OF INVENTION: By Assessing DNA Methylation
 ; FILE REFERENCE: 5013.1012
 ; CURRENT APPLICATION NUMBER: US/10/204,708
 ; PRIOR FILING DATE: 2003-05-06
 ; PRIOR APPLICATION NUMBER: PCT/EP01/03971
 ; PRIOR FILING DATE: 2001-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019058.8
 ; PRIOR FILING DATE: 2000-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019173.8
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: DE 10032529.7
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: DE 10043826.1
 ; PRIOR FILING DATE: 2000-09-01
 ; NUMBER OF SEQ ID NOS: 98
 ; SEQ ID NO 6
 ; LENGTH: 6669
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
 US-10-204-708-6
 Query Match 2.2%; Score 106; DB 4; Length 6669;
 Best Local Similarity 44.0%; Pred. No. 5,9e-11;
 Matches 930; Conservative 0; Mismatches 1156; Indels 31; Gaps 10;

QY 680 ATTGCACTGTCAGACATGCTATATGCTTAAGTGGAGCGTATCATAGAAATATT 739
 DB 6369 AAAACATCTCTCTCCACATATTTATTTATTTCTTAAAAAAATTAATTAATTTATTTCTT 6310
 QY 740 ACACATATTAACACCTTACCGAAAAATTTAAATATTTTACCAATATCATCTACAGAC 799
 DB 6309 AAACATATCAAAATATTTCTTAAAAACCGTACCTTAATTAATTTCTTAATTAATTAATTCAC 6250
 QY 800 AAGATATATGCAACCGACATATAGACGAATTTATTCGACACATATCATATTTATTC 859
 DB 6249 AA-ATTAACAAAAAATTTCTTATCTTAATTAATTAATTAATTAATTAATTAATTAATTA 6191
 QY 860 TTAATATTTCTTACACACATTTACGACCATATATGACAAATACACAAACAGAT 919
 DB 6190 TATTAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6131
 QY 920 ATGTGTAATTAATCTCAATATTTATTTGATTAAGATGTAATTAATTAATTAATTAATTA 979
 DB 6130 ATAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6071
 QY 980 TTTCCAGTGAATATGACAGAAATTTTACAAATATATCAATTTTATCTCTCTGAT 1039
 DB 6070 ATACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6011
 QY 1040 CCAAGTTTGTCTAAGAAATGCAACATCTTCTGATATGATATGATATGATATGATATGAT 1099
 DB 6010 ACATATATTTTAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5951
 QY 1100 AATAAAGACTGTGATCATGATGATGAACAAATAGATAGATGATGATGATGATGATGAT 1159
 DB 5950 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5891
 QY 1160 TACCATCAAGATATTA-TAGGTGTACAGATGAGAGTGAATTAAGATTAATTAATTAATTA 1218
 DB 5890 ATAAATATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5831
 QY 1219 TGATGTGAATATGCTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1278
 DB 5830 AATATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5776
 QY 1279 TAAAGCATGAATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1338
 DB 5775 CTCTTAACCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5716
 QY 1339 TACATTAATTTGTCATGATGCTTGAAGATCTAATTAATTAATTAATTAATTAATTAATTA 1398
 DB 5715 CCTTGTATGCTCTCTACTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5659
 QY 1399 ACGCAATATGCTTTGACATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1458
 DB 5658 CCCAACTCTCTGATGACCTACTACTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5599
 QY 1459 TCACAACTTAAGCTTTGGAAGAAACGATCTGTAATTAATTAATTAATTAATTAATTAATTA 1518
 DB 5598 AAACCTA-AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5541
 QY 1519 AAACTTAATGTAAGTCACTCAATTTGATATCTATTAATTAATTAATTAATTAATTAATTAATTA 1578
 DB 5540 AAAATCTCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5481
 QY 1579 ATACTAGATTAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1638
 DB 5480 ACCAAGCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5421
 QY 1639 ACACATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1698
 DB 5420 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5361
 QY 1699 GCAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1758
 DB 5360 CCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5309

QY 1759 AAGTATTAATTAATTTCTAGCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1818
 DB 5308 AATTTCTACTTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5249
 QY 1819 TAAACATTTGTTACATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1878
 DB 5248 AACCTTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5189
 QY 1879 ATATGTTGCACTTTTACATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1938
 DB 5188 ATAACTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5129
 QY 1939 AACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1998
 DB 5128 AATTCGAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5069
 QY 1999 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2058
 DB 5068 AATCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5009
 QY 2059 TATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2118
 DB 5008 CTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4953
 QY 2119 GGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2178
 DB 4952 AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4893
 QY 2179 AACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2238
 DB 4892 ACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4835
 QY 2239 AGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2298
 DB 4834 ACCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4775
 QY 2299 CAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2358
 DB 4774 TTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4718
 QY 2359 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2418
 DB 4717 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4658
 QY 2419 AAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2478
 DB 4657 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4600
 QY 2479 ATTAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2494
 DB 4599 ACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4584

RESULT 12

US-09-328-352-333

Sequence 333, Application US/09328352

Patent No. 6562958

GENERAL INFORMATION:

APPLICANT: Gary L. Breton et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

FILE REFERENCE: GTC99-03PA

CURRENT APPLICATION NUMBER: US/09/328,352

CURRENT FILING DATE: 1999-06-04

NUMBER OF SEQ ID NOS: 8252

SEQ ID NO 333

LENGTH: 1077

TYPE: DNA

ORGANISM: Acinetobacter baumannii

US-09-328-352-333

Query Match

Best Local Similarity

2.2%; Score 105.8; DB 4; Length 1077;

47.1%; Pred. No. 4.7e-11;

QY 2164 AGCTCAATTAATTAACCAATTTCTTAATTAACCACTACTTAACATAGGAACCAT 2223
DB 7209 AGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7150
QY 2224 AGTATATCTACTAGTATAGATATCCGTACAGCATGTATCAATTAATTAATTAATTCG 2283
DB 7149 TAAATTTAAAAAACAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 7094
QY 2284 TTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2343
DB 7093 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7034
QY 2344 TATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2403
DB 7033 AATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6975
QY 2404 CATCATATCTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2458
DB 6974 AATCATATCTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6920

RESULT 14

US-10-204-708-73/c
Sequence 73, Application US/10204708
Patent No. 6677731
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: RIEBENBROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
FILE REFERENCE: 5013.1012
CURRENT APPLICATION NUMBER: US/10/204,708
CURRENT FILING DATE: 2003-05-06
PRIOR APPLICATION NUMBER: PCT/EP01/03971
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: DE 10019058.8
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: DE 10019173.8
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 98
SEQ ID NO: 73
LENGTH: 5152
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-73

Query Match 2.2%; Score 104; DB 4; Length 5152;
Best Local Similarity 43.9%; Pred. No. 1.3e-10;
Matches 987; Conservative 0; Mismatches 1220; Indels 41; Gaps 11;

QY 273 TCTTACATTAAGAAACACGATGATTCAGAAATACCTCTCTCTCTCTCTCTCTCTCTCTCT 332
DB 4735 TCTTACATTAAGAAACACGATGATTCAGAAATACCTCTCTCTCTCTCTCTCTCTCTCTCT 4676
QY 333 CCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 391
DB 4675 ACAATTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4616
QY 393 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 451
DB 4615 ACTTTTAAAAAACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4556
QY 452 CAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 511
DB 4555 TCACATTAAT--ACTTTAAAAAATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4498

QY 512 ATTAGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 571
DB 4497 AATCTACCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4438
QY 572 TATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 631
DB 4437 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4378
QY 632 ATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 691
DB 4377 ATACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4318
QY 692 GCATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 751
DB 4317 CCGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4258
QY 752 ACTTACCGAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 811
DB 4257 ACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4198
QY 812 ACCGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 871
DB 4197 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4138
QY 872 ACACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 931
DB 4137 CTACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4079
QY 932 ACTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 991
DB 4078 AATCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4019
QY 992 ATAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1051
DB 4018 ATCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3959
QY 1052 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1111
DB 3958 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3899
QY 1112 GGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1171
DB 3898 TTAATTTTA-----AAATTTAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3846
QY 1172 ATTATAGGTACAGATTGAGCTGCATTAAAGAAATTAATTAATTAATTAATTAATTAATTAATTA 1231
DB 3845 ATTAACGAAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3786
QY 1232 ATGCTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1289
DB 3785 TCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3726
QY 1290 ATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1349
DB 3725 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3666
QY 1358 GTCCATCATGCTGTGAGATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1409
DB 3665 TTACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3606
QY 1410 CT-TTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1468
DB 3605 TTAATTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3546
QY 1469 AAGCTTTGAAGAAACGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1528
DB 3545 AAAAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3486
QY 1529 GTAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1588
DB 3485 TTA-ATTACTACTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 3427

QY 1589 AAGCCATATCAACATATTTCTTGTCTATCATCTCTTGTGATTTTCAACCAATGAT 1648
 DB 3426 TAAACATATTTATTAATTAACCAAAATATTTTAAAA-----AATATTTACACCAAAATAT 3373
 QY 1649 AATTAACGGTCAAAATATAGATATACAGCTTACACACATATATATAAAGCAGTCAAT 1708
 DB 3372 TTTAAAAATATTAACATTAATTTTCAAAACATTAATCATTAATACCTTTAAAAATTTAT 3313
 QY 1709 ACCCATGTTAATCACTCACTAAGTATGTAATTAATTAATTAATTAATTAATTAAT 1768
 DB 3312 TTTAAAAAACTATTCACCAACATTTTATTAATTAATTAATTAATTAATTAATTAAT 3253
 QY 1769 TTAATTTCCAGTATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1828
 DB 3252 TATCTTCAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3193
 QY 1829 TTAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1888
 DB 3192 AATATTTCCCAAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3133
 QY 1889 AATATTTTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 1948
 DB 3132 CTAAATCCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3073
 QY 1949 AAGAACCAATATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1999
 DB 3072 CATTAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3013
 QY 2000 TTAATATGATGCTCTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2059
 DB 3012 TAAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2953
 QY 2060 ATGATTAATCTGTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2113
 DB 2952 ATTATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2893
 QY 2114 AAAAAAGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2173
 DB 2892 TTAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2833
 QY 2174 AATTAACCAATTTCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2233
 DB 2832 ACCTCTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2173
 QY 2234 ACCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2293
 DB 2172 AATAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2113
 QY 2294 TTAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2353
 DB 2112 CAATCCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2258
 QY 2354 TCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2413
 DB 2257 TATAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2598
 QY 2414 TTTTAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2473
 DB 2597 TATTTAAACCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2538
 QY 2474 GAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2501
 DB 2537 TAAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2510

RESULT 15
 US-10-204-708-14/c
 ; Sequence 14, Application US/10204708
 ; Patent No. 6677731
 ; GENERAL INFORMATION:
 ; APPLICANT: OLEK, Alexander
 ; APPLICANT: PIENENBROCK, Christian
 ; APPLICANT: BERLIN, Kurt

; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
 ; FILE REFERENCE: 5013.1012
 ; CURRENT APPLICATION NUMBER: US/10/204.708
 ; PRIOR FILING DATE: 2003-05-06
 ; PRIOR APPLICATION NUMBER: PCT/EP01/03971
 ; PRIOR FILING DATE: 2001-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019058.8
 ; PRIOR FILING DATE: 2000-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019173.8
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: DE 10032529.7
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: DE 10043826.1
 ; PRIOR FILING DATE: 2000-09-01
 ; NUMBER OF SEQ ID NOS: 98
 ; SEQ ID NO 14
 ; LENGTH: 6113
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
 ; US-10-204-708-14
 Query Match 2.2%; Score 102.2; DB 4; Length 6113;
 Best Local Similarity 43.9%; Pred. No. 2.9e-10;
 Matches 948; Conservative 0; Mismatches 1163; Indels 48; Gaps 10;
 QY 362 TATTCGATCATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 421
 DB 6062 TATCCAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6003
 QY 422 AATATCTCAATAGCTTTTACCAAGAACTACAAATTAATTAATTAATTAATTAATTAAT 481
 DB 6002 AATATCAATCACTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5943
 QY 482 ATATTTCAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 539
 DB 5942 AAAAAATGCTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5883
 QY 540 TTAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 599
 DB 5882 TAAATTTAAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5823
 QY 600 TGACTGTTTCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 659
 DB 5822 TTAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5768
 QY 660 CATTGGCAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 719
 DB 5767 TAAATCCATTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5708
 QY 720 GAGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 779
 DB 5707 CTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5648
 QY 780 CCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 839
 DB 5647 AAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5586
 QY 840 ACATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 899
 DB 5587 AAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5531
 QY 900 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 959
 DB 5530 AACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5471
 QY 960 TCGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1019
 DB 5470 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5417
 QY 1020 CAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1079

Db 5416 AAAACCTTACTTAATATAAAAAACCAATCCAAAAATTCATATATATAATATAC 5357
 1080 GATATGCTCTGCTGCTGATATATAAGAACTGGGATGATCATGATATATAACATATAGTA 1139
 5356 ATATATAACATATCTCAATATAAAAAATATATATACAAAAACCAATCAATATATCTA 5297
 1140 ATGATGCTTAAAAATATATATACATCAAGATATATATAGTGTGATGATGAGTGCAT 1199
 5296 AATATATAAATAAAAAATATATATACAAAAATATACATATAAAAAATATATCAAAA 5237
 1200 TAAAGAACATATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1239
 5236 TAAAGAACATATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5177
 1260 CTGTTGCTCTGCTGCTGATATATAAGGATGATGATGATGATGATGATGATGATGATG 1319
 5176 TTTATATAAATCT-----ATACCAATTTTAAAAATTTTCTATATTTTAAAAATA 5123
 1320 ATCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1379
 5122 AATCTAAAAAAACCAATATATACAAAACTAAAAAAATAAAAATATATATATATATATAT 5063
 1380 CATGTCACAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1439
 5062 AACACCTACTATATAAAAACTAAAAATTAATAAACTATACCAATATATATATATATA 5003
 1440 ATATATAATCATTAATATATATATATATATATATATATATATATATATATATATATAT 1499
 5002 TATATAAACAATATAAATATCTATATATATATATATATATATATATATATATATATAT 4943
 1500 CACTGATATATATATATATATATATATATATATATATATATATATATATATATATAT 1559
 4942 AATCATATTTCTTAAAAATTAATATATATATATATATATATATATATATATATATATAT 4883
 1560 ATTTATATATATATATATATATATATATATATATATATATATATATATATATATATAT 1619
 4882 AATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 4823
 1620 TCATCTTATATATATATATATATATATATATATATATATATATATATATATATATAT 1679
 4822 CATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 4763
 1680 TTCACCATATATATATATATATATATATATATATATATATATATATATATATATAT 1733
 4762 TAAATATAAACAATATATATATATATATATATATATATATATATATATATATATATAT 4703
 1734 ATGTAATATATATATATATATATATATATATATATATATATATATATATATATATAT 1793
 4702 AATCTATATATATATATATATATATATATATATATATATATATATATATATATATAT 4643
 1794 ACNATATCTATCTCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCT 1853
 4642 CCAAAAAACAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 4583
 1854 TAAATCTTTAC--TATATATATATATATATATATATATATATATATATATATATATAT 1910
 4582 TCTAACCAACGACATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTAT 4523
 1911 ATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 1970
 4522 AATATCAAAACCAAAATTTCAAAACCAATCTATCTATCTATCTATCTATCTATCTATCTAT 4463
 1971 TTTATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2025
 4462 TAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 4403
 2026 --TTTATATATATATATATATATATATATATATATATATATATATATATATATATAT 2083
 4402 CGAATAAATAATCTATATATATATATATATATATATATATATATATATATATATATAT 4343
 2084 ATTAGTTCTTATATATATATATATATATATATATATATATATATATATATATATATAT 2143

Db 4342 ACACTCCAACTTAAAGACAAAAAGAACTCCATCTCAAAATATAACTAAATATAATAT 4283
 2144 ATATGAAGAGATATATATATATATATATATATATATATATATATATATATATATATAT 2203
 4282 AATATATAAATAATATATATATATATATATATATATATATATATATATATATATATAT 4223
 2204 CTCTAAACATGAGAACCAATATATATATATATATATATATATATATATATATATATAT 2263
 4222 TCCATATAAACAATAAACA-----TAAACAAACTAAAAAATAAACAATAAATATATATA 4167
 2264 CAATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2323
 4166 AATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 4107
 2324 AAAAACTCAAGTTCTATATATATATATATATATATATATATATATATATATATATATAT 2383
 4106 CAATCTTCAAAATATATATATATATATATATATATATATATATATATATATATATAT 4053
 2384 AGTTATCAATATATATATATATATATATATATATATATATATATATATATATATATAT 2443
 4052 ACTTCAATATATATATATATATATATATATATATATATATATATATATATATATATAT 3993
 2444 ATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2502
 3992 AATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 3934

RESULT 16
 US-09-328-352-1302/c
 ; Sequence 1302, Application US/09328352
 ; Patent No. 6562958
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary L. Breton et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 ; FILE REFERENCE: GTC99-03PA
 ; CURRENT APPLICATION NUMBER: US/09/328,352
 ; CURRENT FILING DATE: 1999-06-04
 ; NUMBER OF SEQ ID NOS: 8252
 ; SEQ ID NO 1302
 ; LENGTH: 1434
 ; TYPE: DNA
 ; ORGANISM: Acinetobacter baumannii
 US-09-328-352-1302

Query Match 2.1%; Score 101.4; DB 4; Length 1434;
 Best Local Similarity 45.1%; Pred. No. 3.2e-10;
 Matches 516; Conservative 0; Mismatches 611; Indels 18; Gaps 3;

2543 GCAATTCGATATCATCTTACCTCCAAAACGAGATATAGTGTTTTAAATAGTCTTTCA 2602
 1313 GCAATTCGATATATATCTTTTCAATATATTTCTAAATAAAGTTCATAGCTTACCTTGT 1254
 2603 AATTCCTACGAGCTTTACGTAATGCTACAG-----AATTAACCTTACGCTTAAAGCA 2656
 1253 AATTTTTCACGCTTGAAGAGCTGCCACAGTGTGTATGCTTTGAAGACCGAAGCA 1194
 2657 TCATTAATAGCGAATTTAGTACTATATATCTACTGTGTAATCTTTTCTGTATCATCTCC 2716
 1193 AAGGATATGAAATATATGCTCTTGAAGAGCGAGTGAATGAATCATCTATCGTA 1134
 2717 TTAGGAGATTTCAATATTAATATCAATCTATATATATATATATATATATATATATATAT 2776
 1133 GCATCTTCACTTATAGGTAATGCGCCCTCATATATTTCTTATATCGGAACATATACC 1074
 2777 GGCATTCATATATATATATATATATATATATATATATATATATATATATATATATATAT 2836
 1073 GGAAGATGCTGCTGATGAGATATATATATATATATATATATATATATATATATATATAT 1014
 2837 CCAATTTTTCACAGATGCTATATATATATATATATATATATATATATATATATATATATAT 2896
 1013 TCCCAATCCATACAAATTTTCTGATATATATATATATATATATATATATATATATATATAT 954

[illegible]

```

QY      2118 AGGTTAAATTCACATCTCCGCAACAAVATTAAGAGATACATATACAGCGTCAAAVTAATT 2177
Db      4105 TAAAAAAAAAAACCTTACCAATCTTTAAATTATACATAAAAAACCTACACACAAAATTC 4046
QY      2178 AAACAAATTTCTAATATATACACCTACTCTTAAACATAGGAACCAATAGTATATACACT 2237
Db      4045 CATACACAAACCATTAATATATAAACCTCCCTCCATATATACATATAAATATTCCTAATA 3986
QY      2238 AAGTAGATATCCCTACAGCATGTCATCAAAVTAATTAAGAAAAATCGTATATATATATAT 2297
Db      3985 CTTTAAAAAATCAATA-AAAATATATTAACCTTTAAATTAATAAATAAAAAAATTTATTA 3927
QY      2298 ACAATTCAGAAATTTTCCAACTTCACACAAACAAACCTCAAAAGTCTATTATTAACAAATTC 2357
Db      3926 CTAACACTTAAATATACAGATATTAAC----TCTCCCACTAACCTTTAATTAATTAACAATA 3878
QY      2358 GAATATATATACCTCCACCAACATATGTTATCAAAATTCATATATACATCATATCCTTT 2417
Db      3870 AATATTAATATCCACAA-ATTTTAAATTTTAAATATCTATATATTTCTTAAAAAATAT 3812
QY      2418 TAAAGAGTACCTTACATTCATATCAATTAATTAATATA 2456
Db      3811 TATACATATATATTTTAAATTTAATTAATTAATTAATA 3773

RESULT 19
US-10-204-708-32/c
/ Sequence 32, Application US/10204708
/ Patent No. 6677731
/ GENERAL INFORMATION:
/ APPLICANT: OLEX, Alexander
/ APPLICANT: BERLIN, Kurt
/ TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
/ FILE REFERENCE: 5013.1012
/ CURRENT FILING DATE: 2003-05-06
/ PRIOR APPLICATION NUMBER: PCT/EP01/03971
/ PRIOR FILING DATE: 2001-04-06
/ PRIOR APPLICATION NUMBER: DE 10019058.8
/ PRIOR FILING DATE: 2000-04-06
/ PRIOR APPLICATION NUMBER: DE 10019173.8
/ PRIOR FILING DATE: 2000-04-07
/ PRIOR APPLICATION NUMBER: DE 10032529.7
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: DE 10043826.1
/ PRIOR FILING DATE: 2000-09-01
/ NUMBER OF SEQ ID NOS: 98
/ SEQ ID NO 32
/ LENGTH: 8093
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-32

Query Match      2.1%; Score 101; Db 4; Length 8093;
Best Local Similarity 47.1%; Pred. No. 5.1e-10;
Matches 546; Conservative 0; Mismatches 595; Indels 18; Gaps 7;

QY      1512 TAGCAAAAAAAGCTTAATGCTAGACATATACCAATTTGCATATCTATTAACATTTATACAGT 1571
Db      6716 TAACTTAATTCCTTAATTTACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6657
QY      1572 AGAATATATACACTGATTAAGCCATATCAACATATATAT-CTTAGTCTATCATCCTTTGA 1630
Db      6656 AAACAAAAGCTTAAACCAAAAATTTTAAATATATATATATCTTAAAAAATATCTATTAATA 6597
QY      1531 TAATTTCAACAGATGATTAATTAACGCTCAAAAATATAGAAATCAACGTTACACACATA 1690
Db      6596 TAAATTAATTCCTATTAATTAATTTTACATATATAAAGCTTTAAATTTACCTTTTCAATA 6537

```

Page 18

RESULT 20
US-10-204-708-20/c

| | | | | | |
|----|------|-----------------------------------------------------------------|------|-----------------------------------------------------------------|------|
| Db | 1912 | TATATAGAACATAATTTTACATATAAATPACTAATATAAAGAGAACCAATACATATATTAAT | 1974 | TATATAGAACATAATTTTACATATAAATPACTAATATAAAGAGAACCAATACATATATTAAT | 1974 |
| Qy | 1912 | TATATAGAACATAATTTTACATATAAATPACTAATATAAAGAGAACCAATACATATATTAAT | 1974 | TATATAGAACATAATTTTACATATAAATPACTAATATAAAGAGAACCAATACATATATTAAT | 1974 |
| Db | 5864 | TTTCTTTTCGATATTTT--CAAAACCAATTAATTAATCCCAAAATACAAACCCACAA | 5807 | TTTCTTTTCGATATTTT--CAAAACCAATTAATTAATCCCAAAATACAAACCCACAA | 5807 |
| Qy | 5864 | TTTCTTTTCGATATTTT--CAAAACCAATTAATTAATCCCAAAATACAAACCCACAA | 5807 | TTTCTTTTCGATATTTT--CAAAACCAATTAATTAATCCCAAAATACAAACCCACAA | 5807 |
| Db | 1972 | TATATAGAACATAATTTTACATATAAATPACTAATATAAAGAGAACCAATACATATATTAAT | 1974 | TATATAGAACATAATTTTACATATAAATPACTAATATAAAGAGAACCAATACATATATTAAT | 1974 |
| Qy | 1972 | TATATAGAACATAATTTTACATATAAATPACTAATATAAAGAGAACCAATACATATATTAAT | 1974 | TATATAGAACATAATTTTACATATAAATPACTAATATAAAGAGAACCAATACATATATTAAT | 1974 |
| Db | 5984 | TTATATAATTTAAACCAAAATATCAAAATTTTAAATAAATATTAATAAAACCTCAAAATAA | 5925 | TTATATAATTTAAACCAAAATATCAAAATTTTAAATAAATATTAATAAAACCTCAAAATAA | 5925 |
| Qy | 5984 | TTATATAATTTAAACCAAAATATCAAAATTTTAAATAAATATTAATAAAACCTCAAAATAA | 5925 | TTATATAATTTAAACCAAAATATCAAAATTTTAAATAAATATTAATAAAACCTCAAAATAA | 5925 |
| Db | 1652 | ATTAACTCTTACTATTTTAAATAAATAATAGTTGACACCTTTTACACTACATCATGTA | 1911 | ATTAACTCTTACTATTTTAAATAAATAATAGTTGACACCTTTTACACTACATCATGTA | 1911 |
| Qy | 1652 | ATTAACTCTTACTATTTTAAATAAATAATAGTTGACACCTTTTACACTACATCATGTA | 1911 | ATTAACTCTTACTATTTTAAATAAATAATAGTTGACACCTTTTACACTACATCATGTA | 1911 |
| Db | 5924 | ATTATACCTTTCTAATTTTAAATTTATTTTATTTACATATATTTATTTTATTTTAAAT | 5866 | ATTATACCTTTCTAATTTTAAATTTATTTTATTTACATATATTTATTTTATTTTAAAT | 5866 |
| Qy | 5924 | ATTATACCTTTCTAATTTTAAATTTATTTTATTTACATATATTTATTTTATTTTAAAT | 5866 | ATTATACCTTTCTAATTTTAAATTTATTTTATTTACATATATTTATTTTATTTTAAAT | 5866 |
| Db | 2032 | CATCCTTGACATATTTCAAAACATATAAATATGATATAAAGCTGTTTTTATGCCCATTAAGTC | 2091 | CATCCTTGACATATTTCAAAACATATAAATATGATATAAAGCTGTTTTTATGCCCATTAAGTC | 2091 |
| Qy | 2032 | CATCCTTGACATATTTCAAAACATATAAATATGATATAAAGCTGTTTTTATGCCCATTAAGTC | 2091 | CATCCTTGACATATTTCAAAACATATAAATATGATATAAAGCTGTTTTTATGCCCATTAAGTC | 2091 |
| Db | 5746 | TTTATATCTTAATATATCTTAATATAAATATATAAATAAATCTACATATATCTTAATATATA | 5687 | TTTATATCTTAATATATCTTAATATAAATATATAAATAAATCTACATATATCTTAATATATA | 5687 |
| Qy | 5746 | TTTATATCTTAATATATCTTAATATAAATATATAAATAAATCTACATATATCTTAATATATA | 5687 | TTTATATCTTAATATATCTTAATATAAATATATAAATAAATCTACATATATCTTAATATATA | 5687 |

Db 6108 AAAAAAGTAAATTAACCTTAACCTCCCTTAACCTTAATTAATTAATTAACCTAACAA 6049
 QY 1582 ACTAGATAAAGCATATCAACATATTAATCTTGTCTATCACTCTTGTATTTTCAACA 1641
 Db 6048 CCCCTAAAAATAAAAATAAAAATATATATATTAATTAATTAATTAATTAATTAAT 5989
 QY 1642 CATATATTAATTAAGGTCACAAAATATAGATTAACAGCTTACCAACATATTAATTA 1701
 Db 5988 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5929
 QY 1702 GTCAATATCCCATCTGTATTAATCACTAGATATTAATTAATTAATTAATTAATTA 1760
 Db 5928 TATATATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5869
 QY 1761 GTATATCTTAATTAATTTCTTGTATTAATTAATTAATTAATTAATTAATTAATTA 1820
 Db 5868 ATAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5809
 QY 1821 ACCAATTTT-----TATCAGATTATTAATTAATTAATTAATTAATTAATTAAT 1872
 Db 5808 TCCATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5749
 QY 1873 AAAATTAATTAATTTGTGCACT--TTTTCATCACTATCATATTAATTAATTAATTA 1930
 Db 5748 AAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5689
 QY 1931 CATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1990
 Db 5688 TAAACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5629
 QY 1991 TACGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2050
 Db 5628 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5569
 QY 2051 ACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2110
 Db 5568 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5509
 QY 2111 CCCAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2166
 Db 5508 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5449
 QY 2167 TCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2226
 Db 5448 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5389
 QY 2227 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2286
 Db 5388 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5329
 QY 2287 TGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2345
 Db 5328 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5289
 QY 2346 --TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2403
 Db 5288 TGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5209
 QY 2404 CATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2463
 Db 5208 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5149
 QY 2464 TGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2496
 Db 5148 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5116

RESULT 22
 US-09-543-681A-3331
 ; Sequence 3331, Application US/09543681A
 ; Patent No. 6605709
 ; GENERAL INFORMATION:

; APPLICANT: GARY BRETON
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
 ; FILE REFERENCE: 2709.1002-001
 ; CURRENT APPLICATION NUMBER: US/09/543,681A
 ; PRIOR FILING DATE: 2000-04-05
 ; PRIOR APPLICATION NUMBER: US 60/128,706
 ; NUMBER OF SEQ ID NOS: 8344
 ; SEQ ID NO 3331
 ; LENGTH: 1023
 ; TYPE: DNA
 ; ORGANISM: Proteus mirabilis
 ; US-09-543-681A-3331

Query Match 2.0%; Score 95.6; DB 4; Length 1023;
 Best Local Similarity 47.3%; Pred. No. 3.6e-09;
 Matches 463; Conservative 0; Mismatches 494; Indels 21; Gaps 5;

QY 638 ATCAGTATTTTGAAGTGCAGATCATTTGGCAGACGAAATGCAATTCAGCTGCAGACAT 697
 Db 22 ATGACAGTTATGCTGCGGTCATACGCGACCGGTTAGCATTAACCTTAGCGGTAT 81
 QY 698 GGTATATCACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 757
 Db 82 GGGCAGATGTTGCTGCTGGGGGATGATCCAGACAGTGGCGCATTAAGACAGCG 141
 QY 758 CGAAAAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 817
 Db 142 CGCTGATTAACAGCTTTTCTGCTGATGTTCTTCTCGATAGTTATATATGAAGT 201
 QY 818 AATTAAGACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 877
 Db 202 TCTTTCAAAAAGCGATTAAGACGCGCTA--ATATCTTGTGTATCCAGCCAT 258
 QY 878 CAATTCAGCAGCATATGATACAAATTAACACAAACAGATGATGTAATTAATTAATTAAT 937
 Db 259 GTGTTGGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 312
 QY 938 AATTAATTTGATTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 994
 Db 313 GTTGTGGGCGCAAAAGCTGTAAGACATATGATGCTTTATTAAGATGTTGCC 372
 QY 995 GCAAGAAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1054
 Db 373 CGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 432
 QY 1055 GAAATTCAGAAATCTTCTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1114
 Db 433 GAGCTGGCGGGGTTGCCACCGGATGACATGCGCTCCAGCGATATATCTTTTAA 492
 QY 1115 GATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1171
 Db 493 GAGCAGTACACAGCTATTTCAATGTTGTAATTAATTAATTAATTAATTAATTAATTAAT 552
 QY 1172 ATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1231
 Db 553 TTATGCGGTGCAATGCGGTGCGCTTTAAATTAATTAATTAATTAATTAATTAATTAATTA 612
 QY 1232 ATGCGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1291
 Db 613 TGTGATGTTGAGATTTGCGCTTAATGCGCTACGCACTGATTAATTAATTAATTAATTAAT 672
 QY 1292 GAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1351
 Db 673 GAATAGAGCGGTTAGTAAAGCTTAGT-----GCAGATGCGCACTTTTATAGGCG 726
 QY 1352 CATATATGTTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1411
 Db 727 ATGCGTGGTTTGGGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 786
 QY 1412 TTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1471

QY 2876 AATCCGGATATCTGTACATATTTCTACTAAGACGGTACTTATTTGAGACATTT 2935
Db 668 AACTTACTATTTTCTGTTAAACACCTACTTACCAATATTTATTTTACATACAAATTT 609
QY 2936 AATCTATATATA 2948
Db 608 TATTAATAATCGTA 596

RESULT 24
US-10-204-708-59/c
; Sequence 59, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: FIEBENROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 59
; LENGTH: 6156
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-59

Query Match 2.0%; Score 94.6; DB 4; Length 6156;
Best Local Similarity 42.5%; Pred. No. 7.4e-09;
Matches 850; Conservative 0; Mismatches 1134; Indels 15; Gaps 6;

QY 351 ATAGAGTTCTTATTCGACATATTAATTAATCAATCAAAACAAATATATAGTATATC 410
Db 2031 ATATTAATCTTATTAATCTTCTTACAGCTCTTCAAAAATTAATATACCAATTAC 1972

QY 411 TAAATTAATTAATAAATCCCAATAGCTTTAGCAAGACCTACAAATATATAGACACTT 470
Db 1971 TATATTAATAAATATACCTATTAATTAATTAATTAATTAATTAATTAATTAATTA 1912

QY 471 AACTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 530
Db 1911 AAAACCATATCAACACTTCAATATATATATTAATTAATTAATTAATTAATTAATTA 1852

QY 531 CTGACATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 590
Db 1851 AATTAATAAATTTTCTTCAAAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1792

QY 591 TATAATTTTGAATGTTCTTGTATATAATTAATTAATTAATTAATTAATTAATTAAT 650
Db 1791 TATTAATTAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1732

QY 651 GTGCGAGATCATTTGGCAGAGATAGCAATTTGACCTGACGACATGTATATCACTTA 710
Db 1731 ACTCTTTCTAACTTACCAAAAACCTCACTTACCTATATTAATTAATTAATTAATTA 1672

QY 711 ACTTATGGGAGCTGATCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 770
Db 1671 AATATCAACAATATTTTCAACACCTTAACCAATTTTAAATTAATTAATTAATTAAT 1612

QY 771 AATATTTACCAATATATATCTACAGACATATATATGCAACCAATATATAGCAAG 830
Db 1611 TATATATACCAACCCCTATATATATTAATTAATTAATTAATTAATTAATTAATTA 1557

QY 831 TATATCTGACACATATATATATATATATATATATATATATATATATATATATATAT 890
Db 1556 ATTTCTTCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1497

QY 891 TATGTAACAATAATACACAAACAGCATATGTATTAATTAATTAATTAATTAATTAAT 950
Db 1496 AATCCATCTATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1437

QY 951 GTAAGGATATGAAATATATATATATATATATATATATATATATATATATATATAT 1010
Db 1436 ATTAATTAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1377

QY 1011 AATATATCCAAATTTTATATCTCTGTGTCGAAGTTTGTCAAGAAATTAATTAATTA 1070
Db 1376 AATTAATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1323

QY 1071 TTCTGTATATATATATATATATATATATATATATATATATATATATATATATAT 1130
Db 1322 TACCTTATGTAATAATCACTAAATAAATAATTAATTAATTAATTAATTAATTAATTA 1263

QY 1131 CAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 1190
Db 1262 TATATATATATATATATATATATATATATATATATATATATATATATATATATAT 1203

QY 1191 GAGCTGATTAAGAAATATATATATATATATATATATATATATATATATATATATAT 1250
Db 1202 TACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1143

QY 1251 GATATATATATATATATATATATATATATATATATATATATATATATATATATAT 1310
Db 1142 TATATATATATATATATATATATATATATATATATATATATATATATATATATAT 1083

QY 1311 TACGAAAAATATATATATATATATATATATATATATATATATATATATATATATAT 1370
Db 1082 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1023

QY 1371 TAAATATTAATATATATATATATATATATATATATATATATATATATATATATATAT 1430
Db 1022 TAACTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 964

QY 1431 AAGTATGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1490
Db 963 AATATATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 904

QY 1491 CTGTAATACCATATATATATATATATATATATATATATATATATATATATATATAT 1550
Db 903 ATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 844

QY 1551 CTATTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1610
Db 843 ATCTCATATATATATATATATATATATATATATATATATATATATATATATATAT 784

QY 1611 CTATGCTATATATATATATATATATATATATATATATATATATATATATATATAT 1670
Db 783 TATATATATATATATATATATATATATATATATATATATATATATATATATATAT 724

QY 1671 AATATACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1730
Db 723 TATATATATATATATATATATATATATATATATATATATATATATATATATATAT 664

QY 1731 AGATATATATATATATATATATATATATATATATATATATATATATATATATAT 1789
Db 663 TATATATATATATATATATATATATATATATATATATATATATATATATATATAT 604

QY 1790 ATTAACAAATTTCTATATATATATATATATATATATATATATATATATATATATAT 1849
Db 603 ATATTAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 544

QY 1850 AAATTAACCTTTACTATTATTAATAATATGTTGACACTTTTTCACATCATG 1909
DB 543 ACCAAATTTTCTACTTATATATATCTTCAAAATTTTTCATCTCAAAATCACAAT 484
QY 1910 TATTAATGAACATATTATTAATAATTAATAAAGGCAATACATATATTA 1969
DB 483 AAATTAATATATATATATATATATATATATATATATATATATATATATAT 424
QY 1970 -ATTACTAAACCTTACTTACTAGTAATATATATATATATATATATATATAT 2028
DB 423 TCCCTTTATTTTCTATTTTCTATATATATATATATATATATATATATATAT 365
QY 2029 TATCATCTTACATATTTTCAAAATATATATATATATATATATATATATAT 2088
DB 364 TTTATATCTTAAATTTCCCAAAATTCATATACGTAATATATATATATATAT 305
QY 2089 TTCTTAATTAATTTGCTATTTACCCAAAGTTAAATACATCTCCGCAATATG 2148
DB 304 CTCTAATCTTAAACATATATATATATATATATATATATATATATATATAT 245
QY 2149 AAGAAATACATATACAGTCAAAATTAATTAACAAATTTCTAATATACCACTTACTTA 2208
DB 244 TACAATATATATATATATATATATATATATATATATATATATATATATAT 185
QY 2209 AACATAGAACCAATATATATATATATATATATATATATATATATATATAT 2268
DB 184 TATCTTCAAAATTAACAAATCTTCTTAAACATTAACCTTAAACCACTTAA 125
QY 2269 AATTAAGAAATATGTTATGTTATATATATATATATATATATATATATATAT 2328
DB 124 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 65
QY 2329 AACTCAAGTCTTATTA 2347
DB 64 TTTCAAAACATATATATA 46

RESULT 25

US-09-543-681A-1742/C
Sequence 1742, Application US/09543681A
Patent No. 6605709
GENERAL INFORMATION:
APPLICANT: GARY BRETON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
FILE REFERENCE: 2709.1002-001
CURRENT APPLICATION NUMBER: US/09/543.681A
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 1742
LENGTH: 1638
TYPE: DNA
ORGANISM: Proteus mirabilis
US-09-543-681A-1742

Query Match 2.0%; Score 93.8; DB 4; Length 1638;
Best Local Similarity 46.9%; Pred. No. 8.3e-09;
Matches 330; Conservativity 0; Mismatches 367; Indels 6; Gaps 1;

QY 2727 TCATTAATTAATTCATTTATACATTAAGTATTTGTTAGTACCTGGCCATTCAT 2786
DB 1357 TTAATAATACCGCGGCTCATATATTTTCAAGCTCTGTAATTTCCCGCAAGAT 1298
QY 2787 ATGACTGCATTTGCTATTAAGTTCATCACTTATATACATAGTACCAAACTATTTT 2846
DB 1297 AATGGTATATGCTCCAAAGCATGTTGGAATACAGTATATTTTTCATGATCGGC 1238
QY 2847 TACAGATGCTATTCATTAATTAACATCAATTCGGTATATTTGACATATTTCTACTA 2906
DB 1237 GGGCAATTTTTCGTAATAATTTTGGCCATATAGGGATATCTTCAAGTCTGCTCCGTA 1178

QY 2907 AAGAGGTAATCTAATTAAGAGACATTTATATATATATATATATATATATATAT 2966
DB 1177 ATGGCGGATATATATATAGAAAAAATTCAGACATATATAGATCTCTCGAAATTC 1118
QY 2967 CAGCTTTTACTTCACTTCAATATCTTTGGAAGACACAAATATTTCTACATCTATAC 3026
DB 1117 CTGATCCGTCAGTTCATATATATATATATATATATATATATATATATATATAT 1058
QY 3027 TAACAGAAATCTTATATTTTCCCTATATATTTTCCCTCTGTATATATGATATC 3086
DB 1057 GATTAATCTTTATACACCTTATCTTCAATTTTACGCTCTTGAAGAACCTGATATGTT 998
QY 3087 TTAATGCGATGATATGTAATCTGTTACTTATATATATATATATATATATATAT 3146
DB 997 TTGGCTGTAATTCATAGGCGATATGCGCAATTTTCATGAGAAAAAGCTATTTACTG 938
QY 3147 CTGCTCTATATTTCCATATATAGAGAGTACTTATAGACAAATATATATATTTCTCT 3206
DB 937 CCATTTCAATGCAACCCATAGGCGATGTTGACACCG-----TAAAGCGCTTTAT 884
QY 3207 CACTACCAATATATATTAACCAATATATATATATATATATATATATATATATAT 3266
DB 883 CATGCCAAAAAGATCACTTTCTAAGAACCGGAAGTACGTGCAAAATCATATTTT 824
QY 3267 TAAATGAGTATCATACCCCTTGATTTTATATATATATATATATATATATATAT 3326
DB 823 TATGCTGTTATTTTATAGATGACTTAAACGATGATGCTGTCGCAATTAATCTTTAC 764
QY 3327 CAACACCTGGGAACAGATATATATATATATATATATATATATATATATATATAT 3386
DB 763 CAGTACCTGTTTCCGCTAAGTATATATATATATATATATATATATATATATATAT 704
QY 3387 TACTTCTAATTTACGTTAATACAGGAGTATACGACTATTTTC 3429
DB 703 GTTCTAGTACATATATGATGATGATGCTTGTATTAATATCTC 661

RESULT 26

US-10-204-708-57/C
Sequence 57, Application US/10204708
Patent No. 6677731
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIERENBROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
FILE REFERENCE: 5013.1012
CURRENT APPLICATION NUMBER: US/10/204.708
CURRENT FILING DATE: 2003-05-06
PRIOR APPLICATION NUMBER: PCT/EP01/03971
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: DE 10019058.8
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: DE 10019173.8
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 98
SEQ ID NO 57
LENGTH: 6326
TYPE: DNA
ORGANISM: Artificial Sequence
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-57

Query Match 1.9%; Score 92.2; DB 4; Length 6326;
Best Local Similarity 44.9%; Pred. No. 2.1e-08;
Matches 436; Conservativity 0; Mismatches 528; Indels 7; Gaps 2;

1423 AATAGAAAGGTAGAAATATTAATACATTAATAGATACAAACCTTAAGCTTTGAGG 1482
1404 AAAAAACCTCTATATACACATTAATCCCTCTCAAAATCCGAATACACATCCC 3045
1483 AACGAGTACTGTAAACCACTGATATCATTTAGCAAAAAAATGATGAGTACCAT 1542
3044 AACCAATCCACTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2985
1543 TGGCATCTATCTTCAATTTATTAATGAGATATTAATCAATGATTAAGGATACAA 1602
2984 TACTTAACCTTAATCCAAATTAATTAATTAATTAATTAATTAATTAATTAAT 2925
1603 CATATTAATCTATATCATCTCTTGAATTTTCAACACATGATTAATTAAGGTC 1662
2924 CCAAAACCTCACTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2865
1663 AATATGAGATTAACAGTCCACATATTAATTAATTAATTAATTAATTAATTA 1722
2864 AAAAAAATCCAAACCAAAAAAATTAATTAATTAATTAATTAATTAATTAAT 2805
1723 TCCAACTAATGATGATACATTAATTAATTAATTAATTAATTAATTAATTA 1782
2804 AACCAAAATCACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2745
1783 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1842
2744 AAAAAATCTAAACCAAAAAAATTAATTAATTAATTAATTAATTAATTAAT 2690
1843 AATAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1902
2689 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2630
1903 CATCATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1962
2629 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2570
1963 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2022
2569 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2510
2023 AATTCATCATCTGATTAATTAATTAATTAATTAATTAATTAATTAATTA 2082
2509 CCAATTAATCACTATTAATTAATTAATTAATTAATTAATTAATTAATTA 2450
2083 CATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2140
2445 AACCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2390
2141 ACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2200
2389 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2330
2201 CATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2260
2329 AATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2270
2261 TATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2320
2269 TAAACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2210
2321 AACCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2380
2209 ATCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2150
2381 CATAGTATCA 2391
2149 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2139

RESULT 27
US-10-204-708-31/C
; Sequence 31, Application US/10204708
; Patent No. 6577731

GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PLEPENBROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
FILE REFERENCE: 5013.1012
CURRENT APPLICATION NUMBER: US/10/204, 708
PRIOR FILING DATE: 2003-05-06
PRIOR APPLICATION NUMBER: DE 10019058.8
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: DE 10019173.8
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 98
SEQ ID NO 31
LENGTH: 8093
TYPE: DNA
ORGANISM: Artificial Sequence
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-31
Query Match 1.9%; Score 92; DB 4; Length 8093;
Best Local Similarity 45.7%; Pred. No. 2.3e-08;
Matches 811; Conservative 0; Mismatches 890; Indels 75; Gaps 11;
744 AATTAACACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 803
2137 AATTAACACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2078
804 TATATGACACGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 863
2077 AATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2018
864 CATTCCTACACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 923
2017 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1958
924 GTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 983
1957 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1898
984 CCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1039
1897 CCTATCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1838
1040 CCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1099
1837 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1778
1100 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1159
1777 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1718
1160 TACATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1219
1717 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1658
1220 GCAATGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1279
1657 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1607
1280 AATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1339
1606 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1560
1340 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1399


```

Db      3207 TACTTACATTAATAATTAACCTTATACCTTAACTTAACTCTAATTAACGACCACTACT 3148
QY      896 ACACAAATACACACACAAACAGCATATGTGTAAAAATATCTCCAAATATTAATTTGTAGTAAA 955
Db      3147 TTATTTACATTAATAATATACACCAATTAACCAATTAATAATAAATCTTAATAAATATTT--TAAA 3090
QY      956 GGTATCGAAATATACATCTCAATTTCCCGATGAAAATGACAGAAATTTTTCATAT 1015
Db      3089 TCCCAAAAATCTTAATTAATAAATCTTAACCCCTTAATCTTAACCACTCTTAACAACTCT 3030
QY      1016 AATCCAAATTTTAACTCTCTGTCGCAAGTTTGTCTAAGAATTTGAGAACTCTTCTCT 1075
Db      3029 TATAAAATATACCTTCAATTTCAAGTATCTCTAATTTATTAATCTTAATCTTCTCTTAC 2970
QY      1076 TGTATATATATCTGCTGTGTATATTAAGAATTTGTGTATCAATTTGAAACATA 1135
Db      2969 TTTATACATTTTATCCAAATCTCTTAATTTGAAACGCAATTAACCTTAATTAACATACTA 2910
QY      1136 AGTATGATGTTCTTAAATATATATACCATCAAGATATTTAGGTATACAGATTTGAGCT 1195
Db      2909 TTTACTATTTTCTCATATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2856
QY      1196 GCATTAAGAACAATATTTGCAATTTGATGGAATTAATCTGTGGAATAATTTAGTAAAT 1285
Db      2855 ACCCTATATATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2796
QY      1286 AATGCTGTGCTACTGTTATATACCTTAAGCATGATGAATTAATAACCTATATATGA 1315
Db      2795 CCTATATATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2736
QY      1316 AAAATCAATCAATGATCTTCAATCAATTAATTTGCTCATGATCTTGGATCTAATA 1375
Db      2735 TACAAAATCTTAAT--TTCTAATTAATTAACCACTTAACCTAAACCTTTTAAAAATTAAT 2678
QY      1376 TTAACATGTAACAGAAACATTCACGCAATTTGGCTTTGACATGAATTAAGAAAGT 1435
Db      2677 ACATATATTAACAAAGCTCTATTAATTAATAAATTAATTAATTAATTAATTAATTAATTA 2618
QY      1436 AGAAATTA--TAATACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1493
Db      2617 CCCCTTAATCTCAACATTTTAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2558
QY      1494 TAAACGCTGATCATTAAGAAAATACTTAATTTGGA--ACTACCAATTTGCAATTC 1551
Db      2557 AAACCAACCTTAACCAACCAACCAACCTCACTCACTTAACCAATTAATTAATTAAT 2498
QY      1552 TATTTACATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1611
Db      2497 TAACTAAATTAATTAACAGGCGCTACATCCCACTTAATTAATTAATTAATTAATTA 2438
QY      1612 TTAAGCTATCACTCTTGTATTAATTTGACACATGATTAATAAAGGTCAAAATTAATGA 1671
Db      2437 ATTACTTA--AACCTTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2380
QY      1672 ATACAACTTCAACCATATTAATAAAGCAGTCAATTAATTAATTAATTAATTAATTAAT 1731
Db      2379 ACCTAAATTAACAAATTAACCACTTAATTAATTAATTAATTAATTAATTAATTAATTA 2320
QY      1732 GATATGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1791
Db      2319 CATATCAATTTTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2260
QY      1792 AAACAAATTTCTATCTCTATCTCAATTAACCAATTTGATTAATTAATTAATTAATTAAT 1851
Db      2259 AAACAACACCATTAACCTTAACCCGACAAATTAATTAATTAATTAATTAATTAATTAAT 2200
QY      1852 ATTAACCTTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1911
Db      2199 AAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2140
QY      1912 TATATGAACATTAATTT--TTACATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1966
Db      2139 TTTTAAAAATCTATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2080

```

```

QY      1967 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2026
Db      2079 TAAAAAGAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2020
QY      2027 TGTATCACTCTGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2086
Db      2019 ACATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1960
QY      2087 AGTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2146
Db      1959 TATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 1900
QY      2147 TGAAGAAGTATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2205
Db      1899 TAAAAAATTTTATTAACA--AGTTCAATTAATTAATTAATTAATTAATTAATTAATTAAT 1842
QY      2206 CTAAACATTAAGAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2265
Db      1841 ATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1782
QY      2266 AATTAATTAAGAAAAATCTGATATATTAATTAATTAATTAATTAATTAATTAATTAATTA 2325
Db      1781 TATTAACATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1722
QY      2326 CAATACTCAAGTTCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2383
Db      1721 CTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1662
QY      2384 AGTTATCAATTTCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2433
Db      1661 TACTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1612

RESULT 29
US-09-328-352-3585/c
; Sequence 3585, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; SEQ ID NO 3585
; LENGTH: 1293
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-3585

Query Match      1.9%; Score 90.4; DB 4; Length 1293;
Best Local Similarity 46.2%; Pred. No. 3.4e-08;
Matches 341; Conservative 0; Mismatches 391; Indels 6; Gaps 1;

```

DB 888 GAAATTTGGAATTCAGCTAGACGGTAGAGTAGATCATGACGGAAGCCAGCGCTAGC 829
QY 2977 TTCACTTTCAATATCTTTGGAAGAAGACACATATATCTCATCTATCTAACAAGAAAT 3036
DB 828 CTCCTCGGTAAGGTTATATGTTGCTGTATATATTCGACATATGGCTTGAATCGGTTT 769
QY 3037 CTACTATTTTCCCTATATATATTTTCCCTCCTGTATATATCTGATATCTTAATTCGT 3096
DB 768 TTCAACCAACCACTGGCGTAATATCACTGTTTCAACACCGGTATATATTTAGCTTGAAG 709
QY 3097 ATCGATGTAATCTGTACTATCTATCTATATATATATAGTACCAAGATTTGGTCTAT 3156
DB 708 CTCCTAAGAAATCTCGCAATTTCTATCCAGAAACAGGCTTCGCTCGGACGCTTAAA 649
QY 3157 AATTCGAATATGAGAGGTAAGTCTATGAGACATATATATTTACTTTCTCACTACCAAA 3216
DB 648 CACGCTTTAT-----GGTCGCGAATTTGCTCCGTAAGACCTTTTCACTGCTTGAAG 595
QY 3217 TATATTAACAAGATATATATATAGTCTGATAGATAGATAGATATATATATATGAGT 3276
DB 594 GAGTTCACCTTTCTAATATATATATATATTAAGCTGCAAGTTAATCCCAAAAATCTT 535
QY 3277 ATCATACCCCTTGATTTTATGATATATAGCTGACCACTATCTTCTTCCAAACCTTG 3336
DB 534 ATTTTACGTTGGCTATATATATATATATATGCTTTTGCCAAAGTCTTTACCGCTACAGA 475
QY 3337 CGAACCAATATATATATATATATATATATATATATATATATATATATATATATAT 3396
DB 474 TTCCGCGCGGCAAGCACTGSAATCTAGTATATGCACTTTGTAATATATGAAACAT 415
QY 3397 ATTACGTATTAACAGGGA 3414
DB 414 ACGAGTCAATTTCTGATG 397

RESULT 30
US-10-204-708-50/c
Sequence 50, Application US/10204708
Patent No. 6677731
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: BERLIN, Kurt
APPLICANT: PISPENBROCK, Christian
TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
FILE REFERENCE: 5013.1012
CURRENT APPLICATION NUMBER: US/10/204.708
CURRENT FILING DATE: 2003-05-06
PRIOR APPLICATION NUMBER: PCT/EP01/03971
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: DE 10019058.8
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: DE 10019173.8
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 98
SEQ ID NO 50
LENGTH: 6306
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-50

Query Match 1.9%; Score 89.8; DB 4; Length 6306;
Best Local Similarity 44.3%; Pred. No. 5.7e-08;
Matches 583; Conservative 0; Mismatches 722; Indels 10; Gaps 5;
QY 1148 CTAATAATATATACATCAAGATTTATAGTGTACAGATGGAGCTGATTAAGAAC 1207

DB 6212 CTATTAACCTCTTTCACATATATATATATATATATATATATATATATATATATAT 6153
QY 1208 ATATATGCAATTCATATGATATATATATATATATATATATATATATATATATATAT 1267
DB 6152 ACAAATAACCCATCTTACTTAAACATATTTCTATTCATAATATATATATATATAT 6093
QY 1268 ACTGTTATATATATATATATATATATATATATATATATATATATATATATATAT 1327
DB 6092 ATCTTAAACAAAATATATATATATATATATATATATATATATATATATATATAT 6036
QY 1328 ATATATCTTACATATATATATATATATATATATATATATATATATATATATATAT 1387
DB 6035 AACTATATATATATATATATATATATATATATATATATATATATATATATATAT 5976
QY 1388 ACGAATATCAAGCAATATATATATATATATATATATATATATATATATATATAT 1446
DB 5975 AATAACATTTCTATATATATATATATATATATATATATATATATATATATATAT 5916
QY 1447 TACATTAATATATATATATATATATATATATATATATATATATATATATATAT 1506
DB 5915 TTCAATATATATATATATATATATATATATATATATATATATATATATATATAT 5856
QY 1507 ATCATTAAGCAAAAATATATATATATATATATATATATATATATATATATATAT 1566
DB 5855 TATATATATATATATATATATATATATATATATATATATATATATATATATAT 5796
QY 1567 ACATGAGATATATATATATATATATATATATATATATATATATATATATATAT 1626
DB 5795 ATATATATATATATATATATATATATATATATATATATATATATATATATAT 5736
QY 1627 TGATATATATATATATATATATATATATATATATATATATATATATATATATAT 1686
DB 5735 ATTTATATATATATATATATATATATATATATATATATATATATATATATATAT 5676
QY 1687 CATATATATATATATATATATATATATATATATATATATATATATATATATAT 1746
DB 5675 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 5619
QY 1747 CATACAAATATATATATATATATATATATATATATATATATATATATATATAT 1806
DB 5618 AAAAAATATATATATATATATATATATATATATATATATATATATATATATAT 5559
QY 1807 TCCATATATATATATATATATATATATATATATATATATATATATATATATAT 1866
DB 5558 AATCCATATATATATATATATATATATATATATATATATATATATATATATAT 5499
QY 1867 ATTTATATATATATATATATATATATATATATATATATATATATATATATATAT 1925
DB 5498 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 5439
QY 1926 TTTTATATATATATATATATATATATATATATATATATATATATATATATATAT 1985
DB 5438 TTAATAAATCTTATCTTAAACATATATATATATATATATATATATATATATATAT 5379
QY 1986 TTTTATATATATATATATATATATATATATATATATATATATATATATATATAT 2045
DB 5378 TTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 5319
QY 2046 TTCAATATATATATATATATATATATATATATATATATATATATATATATATAT 2105
DB 5318 CAAAAATATATATATATATATATATATATATATATATATATATATATATATATATAT 5259
QY 2106 AATTATATATATATATATATATATATATATATATATATATATATATATATATATAT 2165
DB 5258 AATATATATATATATATATATATATATATATATATATATATATATATATATATAT 5199
QY 2166 GTCAATATATATATATATATATATATATATATATATATATATATATATATATATAT 2223
DB 5198 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 5139
QY 2224 AGTATATATATATATATATATATATATATATATATATATATATATATATATATAT 2283

Db 5138 AATATAAACTCTTAAAAATCAATAAAAAATCAAAAAAATTAACAT 5079
Qy 2284 TATGTATATTTATACAAATTAACAAATTTCCAACTTCAACAAACTCAAGTTCTAT 2343
Db 5078 AATATACCAATTAATCAATCAAAAAATATTCAACTTTTCAATATCAAAAAATCAAAAT 5019
Qy 2344 TATAAATCAATCAAGATATATATATATCTTCAACCAACATGTATCAATCTATATA 2403
Db 5018 TAAAAATCAACCAAAAAAACAATTCATCTTCAATCAATTAATAAAAAAATCCCAAAAT 4959
Qy 2404 CATCATCTCTTTTAAAAAGATGACTTCAATTCATATCAATTAATATATACA 2458
Db 4958 TACCAAAAAATATACATATTAATCAATCTTATATCTTCAATCAAAATTAATAA 4904

RESULT 31
US-10-204-708-10/c
Sequence 10, Application US/10204708
Patent No. 6677731
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENBROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
FILE REFERENCE: 5013.1012
CURRENT APPLICATION NUMBER: US/10/204,708
CURRENT FILING DATE: 2003-05-06
PRIOR APPLICATION NUMBER: PCT/EP01/03971
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: DE 10019058.8
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: DE 10019173.8
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 98
SEQ ID NO 10
LENGTH: 6070
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-10

Query Match 1.9%; Score 89.6; DB 4; Length 6070;
Best Local Similarity 44.2%; Pred. No. 6.2e-08;
Matches 612; Conservative 0; Mismatches 749; Indels 23; Gaps 5;

Qy 1130 ACAATAGATGATGTTTAAATAATATATCCATCAAGATATATAGGTGACAGTT 1189
Db 2957 ACTAATATAAATCCGCTCTCAAAAAAATTAATAAAAAAATTAATTAATTTTCAT 2898
Qy 1190 GGAGCTGATTAAGACATTAATTCGAATTCATGATGGAATGATGCTGGAATAATTA 1249
Db 2897 AAAATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2838
Qy 1250 GGTATTAATGCTGTTCTACTGTATTAATTAAGGATGATGAAATTAACATATAT 1309
Db 2837 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2778
Qy 1310 ATAGCAAAAAATCATCATATAGATCTTCATATTAATTAATTAATTAATTAATTAATTAAT 1369
Db 2777 AACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2718
Qy 1370 CTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1429
Db 2717 TATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2658
Qy 1430 AAAGGTAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1489

Db 2657 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2604
Qy 1490 ACTGTAAAAACCATGATATCATTAAGCAAAAAAATTAATGTAAGAACTACCAATTTGCATA 1549
Db 2603 TCCCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2544
Qy 1550 TCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1609
Db 2543 TTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2484
Qy 1610 TCTTATGATCAATCCCTTGATTAATTTCAACAAATGATTAATTAATTAATTAATTAATTAAT 1669
Db 2483 TCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2424
Qy 1670 GATTAACAGTTTCAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1729
Db 2423 TAAAAATCCCATTTTCAATCTTTTACAACTAAAAAATTAATTAATTAATTAATTAATTAAT 2364
Qy 1730 -----AAGTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1781
Db 2363 ATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2304
Qy 1782 CTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1841
Db 2303 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2244
Qy 1842 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1901
Db 2243 AAAAA--TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2187
Qy 1902 ACATCATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1960
Db 2186 ATAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2127
Qy 1961 ATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2020
Db 2126 ATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2067
Qy 2021 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2080
Db 2066 AATACCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2007
Qy 2081 CCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2140
Db 2006 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1947
Qy 2141 ACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2195
Db 1946 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1887
Qy 2196 ACCACCTACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2255
Db 1886 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1827
Qy 2256 GCATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2315
Db 1826 CTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1767
Qy 2316 ACTTAACAAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2375
Db 1766 ACACACACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1707
Qy 2376 CAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2435
Db 1706 AAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1647
Qy 2436 ATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2495
Db 1646 TATCAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1587
Qy 2496 TTTT 2499
Db 1586 ATAT 1583

QY 1790 ATAAACAATTCCTATCTCTATCTCA-ATTACCAATGTATACAGATTATTAAGT 1848
 Db 2048 AAAAAAAAAATTATATAAATAATCCACAAATCCCTAATTATTAACCTTTAAATTCAA 1889
 QY 1849 CAATTAACCTTTACTATTATTAATAATATATGTGACACCTTTTCTACATCAT 1908
 Db 1988 AATATAATACATTTTATTTAATAAATAATATCTTCAAAACAAATATCTAAATATC 1929
 QY 1908 GTATATATGAACATTAATTTTACATATAATATATAAAGAACCAATCATATTTA 1968
 Db 1928 ACCCTAAATATAAATAAACAATAAATAATATAAACAATAAATAAATCATAT 1869
 QY 1969 AATTAACAACACCTATTACTACATATATACATATGATGCTGACATATAATATTC 2028
 Db 1868 AATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1809
 QY 2029 TATCATCTTGACATATTTCAACATTAATATGATAAAGTGTGTTTATGCCCATTAG 2088
 Db 1808 TCAAAATATTT--AAATACAAACATTAATTTTATAC--TTTATCTCTCAAAATAT 1754
 QY 2089 TCTTAATTAATTTGCTATTTACCCAAAAGTTAAATACACATTCCTGCAACATATG 2148
 Db 1753 AATTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1686
 QY 2149 AAGAAATACATACACGTCAAATTAATTAACAAATTTCTAATATTAACAACCTACTTA 2208
 Db 1695 -AAATATATATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1637
 QY 2209 AACATATGAACCAATATGATATACATGATATGATATGATATGATATGATATGAT 2268
 Db 1636 AATATAATATTAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1577
 QY 2269 AATTAAGAAAAATCGTATATGATATTTATATACATTAACAAATTTCCAACTTCAACAA 2328
 Db 1576 TAACCTAAATATACATATATATTAATAAATAAATAAATAAATAAATAAATAAATA 1517
 QY 2329 AACTCAAGTCTATTTATTAACAATTCAGAAATTTATATCTTCCACAAACATATGTA 2388
 Db 1516 AAAAACTATATATATTTATTAATAAATAAATAAATAAATAAATAAATAAATAAATA 1457
 QY 2389 TCAAAATTCATATATCATCATATCTTTTAAAGATGATGCTCAATTCATCATTA 2448
 Db 1456 TAAATAATATACGTCACAAATTTATTAATTCATATTAATAAATAAATAAATAAATA 1397
 QY 2449 TAATTAATACATGCAATGTTTATTCGA 2475
 Db 1396 TTTCTCTCTTATACATTTTTCAAA 1370

RESULT 33
 US-10-204-708-12/c
 ; Sequence 12, Application US/10204708
 ; Patent No. 6677731
 ; GENERAL INFORMATION:
 ; APPLICANT: OLEK, Alexander
 ; APPLICANT: PIEPENBROCK, Christian
 ; APPLICANT: BERLIN, Kurt
 ; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
 ; TITLE OF INVENTION: by Assessing DNA Methylation
 ; FILE REFERENCE: 5013.1012
 ; CURRENT APPLICATION NUMBER: US/10/204,708
 ; PRIORITY FILING DATE: 2003-05-06
 ; PRIOR APPLICATION NUMBER: PCT/EP01/03971
 ; PRIOR FILING DATE: 2001-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019058.8
 ; PRIOR FILING DATE: 2000-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019173.8
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: DE 10032529.7
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: DE 10043826.1
 ; PRIOR FILING DATE: 2000-09-01
 ; NUMBER OF SEQ ID NOS: 98

; SEQ ID NO 12
 ; LENGTH: 6317
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
 US-10-204-708-12

Query Match 1.9%; Score 88.2; DB 4; Length 6317;
 Best Local Similarity 43.5%; Pred. No. 1.1e-07;
 Matches 925; Conservative 0; Mismatches 1163; Indels 39; Gaps 10;

QY 2295 TCAAGAAATACCTCTCTCTGTTCCATCCATATTTTCCATTTTATACATATCATG 354
 Db 3714 TTCAAAATTCATCCATATATATATATCAATTAATTTCTTTATATATAAATAATATAC 3655
 QY 355 AGTCTTTATTCGACATCATATTAATTAATCAATCAAAACAAATAATATGATCTATCTAT 414
 Db 3654 TACATTAATATCATTTATTTATTCACATCAACAAACAAATACATATTAACATTA 3595
 QY 415 TATATAAATAATCCGATATGCTTTAGCAGAACTACAAATATATATGAGACCTTA--A 472
 Db 3594 ATTAATAATTAATACAACTATATTTTAATATTAATTCATATTAATTTTATATAC 3535
 QY 473 CTATTAATATATTTCAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 532
 Db 3534 ACATTAATTAATAAACAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 3475
 QY 533 GACATATTAACAATATGATATACATTTTCTTAACATGATATATATATATATATATAT 592
 Db 3474 AATAATCTAAACAAATATTTCTTTAAATATTTCAAAATTTCAAAATTTTAATATAT 3415
 QY 593 TAATTTTGAATGTTTCTGATATATTAATTAATTAATTAATTAATTAATTAATTAAT 652
 Db 3414 TACCTAAACCAATTAATAACGAAAAATTTTCTCTCTCTTCAACCAATTTCTTA 3355
 QY 653 GCAGATCATTTGGCAGACCAATGCAATTTGACAGCTGACGACAT----GGTATATG 707
 Db 3354 ATGCTTATATATCCATATACATATTAATTTATATATATTAATTAATTAATTAATTA 3295
 QY 708 TTAATTAATGGGACGTGATATAGAAATATTAACATATTAACATTAACCAAAAAAT 767
 Db 3294 TATATTAATACACATACACACACACCCCTTAATTAATTAATTAATTAATTAATTA 3235
 QY 768 TAAATAATTTACCAATATATCATCTACGAGAAAGATATATGCAACGCAATATATAG 827
 Db 3234 TTCAAAATTAATCTTAACAAT-----TCATAAAAATTAATCTTTATATATTTTAAAA 3181
 QY 828 AAGTATTAATCTGCAACAATATACATATATATCTTAATCTTACATCAACAATTAACGA 887
 Db 3180 AATTTATTAACGACAACTTAATTTTAAATAAATATATATAGCCGCCATATATTAACA 3121
 QY 888 CCAATATGACAAATATACAAACAAACAGCATATGTTAAATAATCTCCATATATTAAT 947
 Db 3120 TCTATTTTAAATAATCTTAACCAATTAATTAATTAATTAATTAATTAATTAATTA 3061
 QY 948 GTAGTAAGGATATCAAAATTAATCACTCACTCAAAATTTCCAGTAATAGCAGAAATTT 1007
 Db 3060 AATTAATAAGTTTCTCTAATAAATAAATAATCCGACAACTATTTTATTAATAATCTAC 3001
 QY 1008 TACATAATTAATCAATTTTAAATCTCTGTCGCAAGTTTGTCTAAAGAAATTTGCGAAC 1067
 Db 3000 TCTATTTTCTTTAAACAATTAATCTCTACAGCCCTCTTAACCTTCTTACCAACAAC 2941
 QY 1068 ATCTTCTGATATAGTACTGCTGATATTAAGAACTTGTGATGATATGATAG 1127
 Db 2940 ACAAAC-----ATTACATTCCAAATTTCCAAATTAATATTAATTTTCCAAATA 2885
 QY 1128 AATCAATTAATGATATGTTCTTAATAATATATATATATATATATATATATATATAT 1187
 Db 2884 AATCAAAATAATATATATTTTCAATATTAATTAATTAATTAATTAATTAATTAATTA 2825
 QY 1188 TTGAGCTGATTAAGAATATATGCAATTTGCAATGATGAAATTAATCGCTGAAAAAAT 1247

[illegible]

| Db | 1750 | AAAAAAATCTCTTCAAAAAAAAAAATTCATTTAAAAAATTCGCTCAAAATTTCTACTC | 1691 |
|----|------|------------------------------------------------------------------|------|
| QY | 2312 | TCCAACTTCAACAACAAACCTCAAGTCTTATTTAAACAATTCAGATATATATACCT | 2371 |
| Db | 1690 | TCATCCTTCCAAATATATACTTATATATTCCTTAATATAAAGATATATAAAAAATTTACATTTA | 1631 |
| QY | 2372 | TCACCAACACATAGTTATCAAAATTTCTA | 2398 |
| Db | 1630 | ACGCACAAAAAGATTTCCCAAAATTCGTA | 1604 |

RESULT 34
US-10-204-708-36/c

; Sequence 36, Application US/102047088
; Patent No. 6677731

GENERAL INFORMATION:

APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; ADDRESSE: PIEDIN Kuch

; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
 ; TITLE OF INVENTION: by Assessing DNA Methylation
 ;

FILE REFERENCE: 5013.1012

```

; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; CURRENT PRIORITY NUMBER: US/10/204,708

```

PRIOR APPLICATION NUMBER: PCI/EPOL/039/1

; PRIOR FILING DATE: 2001-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019058.1
 ; PRIOR FILING DATE: 2000-04-06

PRIOR APPLICATION NUMBER: DE 10019173.1

PRIOR APPLICATION NUMBER: DE 10032529

PRIOR FILING DATE: 2000-06-30

PRIOR FILING DATE: 2000-09-01

NUMBER OF SEQ ID NOS: 98

```

; SEQ ID NO 36
;
; LENGTH: 9347

```

TYPE: DNA

```

;
ORGANISM: Artificial Sequence
FEATURE:
;

```

OTHER INFORMATION: chemically

US-10-204-708-36

| Query Match | 1.98; |
|-------------|-------|
| 1.98; | 1.98; |

Best Local Similarity 42.6%;
Matches 918; Conservative

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1

336 ATTATTACAAATACATAG

DB 8/34 AALACGAAAHAAHCH1IAO

396 AATAAGTACTATCTAAT

Db 8694 CAAAAACAAAATCTATATA

456 TATAATGAGACCTTAAC

Db 8634 CCTAAAAAATAAAAAAT

QY 516 AGCATAATATCACAATCTG

Db 8574 CCTAATATATAAAATAT

10

576 ATACAATAATATAT

Db 8514 ATATAAAATACTTTAAAV

636 AATCAGTATTTAGGTGC

0454 TTTTAAATTTT

DB 8454 TCACACCTAATAATCCAAAC

[illegible]

100

Db 8394 ATTCCAAACAAACCTAAACAAACAAACCTTATCTCTATTAATAAATAATAT 8335
Qy 756 ACCGAAAAATTTAAATATTTTACCACATATCATCTACAGACAAATATGCAACA 815
Db 8334 AACCAAAATATGATATATATATCCCAACATCTAAACAAACGAAATATTAATAC 8275
Qy 816 GCAATATGAGCAATATATATCTGACAAATATATATATATATATATATATATAT 875
Db 8274 CTAATATCAAAATCTTAAACCAACCTTACCAATATATATATATATATATATAT 8215
Qy 876 AACAAATGACCAATATATGACAAATATCAACAAACAGCAATAT---GTGTAAT 931
Db 8214 ATACAAATCTAATATATATATATATATATATATATATATATATATATATAT 8155
Qy 932 ACTCCAAATATATATATATATATATATATATATATATATATATATATATATAT 991
Db 8154 AACATTAATATGCTTCAATCTTAAACAAACAAATATATATATATATATATATAT 8095
Qy 992 ATAGCAGAAATTTTAAATATATATATATATATATATATATATATATATATATAT 1051
Db 8094 ACATCCAACTAAACAAACAAATATATATATATATATATATATATATATATATAT 8039
Qy 1052 AAGAAATTTGACGACATCTTCTGATGATATATATATATATATATATATATATAT 1111
Db 8038 AAAACAAACAAACAAACAAACAAATATATATATATATATATATATATATATATAT 7983
Qy 1112 GGTGATCATGATATATATATATATATATATATATATATATATATATATATATAT 1171
Db 7982 CTTAAATATATATATATATATATATATATATATATATATATATATATATATAT 7923
Qy 1172 ATTAATAGGTATATATATATATATATATATATATATATATATATATATATATAT 1231
Db 7922 AATCACCACATCTACATCTTAAACAAACAAATATATATATATATATATATATAT 7863
Qy 1232 ATCGTGAAAAATTTAGTATATATATATATATATATATATATATATATATATATAT 1291
Db 7862 AATATATATATCTTCAACAAATTTTAAATATATATATATATATATATATATATAT 7803
Qy 1292 GAAATTTAAACCTATATATATATATATATATATATATATATATATATATATATAT 1351
Db 7802 ACTAAATCTTAACTATATATATATATATATATATATATATATATATATATATATAT 7743
Qy 1352 CCATCATGCTTGAGATCTATATATATATATATATATATATATATATATATATATAT 1411
Db 7742 TAAATATATATATCTCTATATATATATATATATATATATATATATATATATATAT 7683
Qy 1412 TTGACATGAAATATGAAATATGAAATATATATATATATATATATATATATATAT 1471
Db 7682 CCACCTCTTTAAATATATATATATATATATATATATATATATATATATATATATAT 7623
Qy 1472 CTGTTGAAAGAACAGTATGTAATATATATATATATATATATATATATATATATAT 1531
Db 7622 ACTTTTAAATATATCTCTCTTCAAAATATATATATATATATATATATATATATAT 7563
Qy 1532 GAATCAATTTGATATATATATATATATATATATATATATATATATATATATATAT 1591
Db 7562 TATTTCTTTTCTCTTTTCTCTTTTCTCTTTTCTCTTTTCTCTTTTCTCTTTTCT 7503
Qy 1592 GGCATATCAAAATATATATATATATATATATATATATATATATATATATATATAT 1651
Db 7502 TTTTAAATTTTATATATATATATATATATATATATATATATATATATATATATAT 7443
Qy 1652 TAAACGGTCAAAATATATATATATATATATATATATATATATATATATATATATAT 1711
Db 7442 TCTAATATATATATATATATATATATATATATATATATATATATATATATATAT 7385
Qy 1712 CATGCTTAACTCAATATATATATATATATATATATATATATATATATATATATATAT 1771
Db 7384 ATTAATGCAAAATATATATATATATATATATATATATATATATATATATATATATAT 7325
Qy 1772 AATTTCTAGCTATATATATATATATATATATATATATATATATATATATATATATAT 1831
Db 7324 TAACTTAATATATATATATATATATATATATATATATATATATATATATATATATAT 7267

Qy 1832 TCAGATTAATTAATAGTCAATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1891
Db 7266 TTTTATCTCAACTATATATATATATATATATATATATATATATATATATATATATATAT 7207
Qy 1892 CTTTTCATCATCATGATATATATATATATATATATATATATATATATATATATATATAT 1951
Db 7206 ATTACTCTCAATCAAAATATATATATATATATATATATATATATATATATATATATAT 7150
Qy 1952 AACCAATCATATATATATATATATATATATATATATATATATATATATATATATATAT 2011
Db 7149 AATTAATCAAAATATATATATATATATATATATATATATATATATATATATATATAT 7090
Qy 2012 CCGACATATATATATATATATATATATATATATATATATATATATATATATATATAT 2071
Db 7089 TCACATATCAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7030
Qy 2072 TTTTATATGCCATATATATATATATATATATATATATATATATATATATATATATAT 2131
Db 7029 CTAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 6971
Qy 2132 ATTCCGACACATATATATATATATATATATATATATATATATATATATATATATATAT 2189
Db 6970 CTAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 6911
Qy 2190 AATATATCACTCTATATATATATATATATATATATATATATATATATATATATATAT 2245
Db 6910 AATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 6851
Qy 2250 CCGACATATATATATATATATATATATATATATATATATATATATATATATATATAT 2309
Db 6850 CCATATCAAAATATATATATATATATATATATATATATATATATATATATATATATAT 6791
Qy 2310 TTCCATCTTCAACAAACAAACCTAAAGTCTATATATATATATATATATATATATATAT 2369
Db 6790 AAAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6731
Qy 2370 CTTACACATATATATATATATATATATATATATATATATATATATATATATATATAT 2429
Db 6730 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6671
Qy 2430 TTCAATATCAATATATATATATATATATATATATATATATATATATATATATATATAT 2482
Db 6670 TATTAACCCCAAAATATATATATATATATATATATATATATATATATATATATATATAT 6618

RESULT 35
US-10-204-708-35/c
Sequence 35, Application US/10204708
Patent No. 6677731
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENBROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
FILE REFERENCE: 5013.1012
CURRENT APPLICATION NUMBER: US/10/204,708
PRIORITY FILING DATE: 2003-05-06
PRIORITY FILING DATE: PCT/EP01/03971
PRIORITY FILING DATE: 2001-04-06
PRIORITY FILING DATE: 2001-04-06
PRIORITY FILING DATE: 2000-04-06
PRIORITY FILING DATE: 2000-04-07
PRIORITY FILING DATE: 2000-04-07
PRIORITY FILING DATE: 2000-06-30
PRIORITY FILING DATE: 2000-06-30
PRIORITY FILING DATE: 2000-09-01
PRIORITY FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 98
SEQ ID NO 35
LENGTH: 9347
TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:
OTHER INFORMATION: Chemically treated genomic DNA (Homo sapiens)
US-10-204-708-35Query Match 1.8%; Score 86.8; DB 4; Length 9347;
Best Local Similarity 43.1%; Pred. No. 2.2e-07;
Matches 597; Conservative 0; Mismatches 767; Indels 20; Gaps 3;

1150 AAAATATATATACCAATGATATTATAGTACAGATGGAGCTGCAATTAAGAACAT 1209
DB AAAATATATATACCAATGATATTATAGTACAGATGGAGCTGCAATTAAGAACAT 1209
9335 AAAATATATATACCAATGATATTATAGTACAGATGGAGCTGCAATTAAGAACAT 9276
1210 AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1269
DB AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1269
9275 AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9216
DB AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9216
1270 TGTATATATATACCAATGATATTATAGTACAGATGGAGCTGCAATTAAGAACAT 1329
DB AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1329
9215 AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9156
DB AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9156
1330 AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1389
DB AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1389
9155 AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9096
DB AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9096
1390 AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1449
DB AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1449
9095 AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9036
DB AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 9036
1450 AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1509
DB AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1509
9035 AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8976
DB AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8976
1510 AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1569
DB AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1569
8975 AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8916
DB AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8916
1570 TGAATATATATACCAATGATATTATAGTACAGATGGAGCTGCAATTAAGAACAT 1629
DB TGAATATATATACCAATGATATTATAGTACAGATGGAGCTGCAATTAAGAACAT 1629
8915 TGAATATATATACCAATGATATTATAGTACAGATGGAGCTGCAATTAAGAACAT 8856
DB TGAATATATATACCAATGATATTATAGTACAGATGGAGCTGCAATTAAGAACAT 8856
1630 AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1689
DB AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1689
8855 AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8798
DB AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8798
1690 AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1749
DB AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1749
8797 AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8738
DB AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8738
1750 AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1809
DB AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1809
8737 AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8678
DB AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8678
1810 AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1869
DB AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1869
8677 AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8618
DB AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8618
1870 TATATATATATACCAATGATATTATAGTACAGATGGAGCTGCAATTAAGAACAT 1929
DB TATATATATATACCAATGATATTATAGTACAGATGGAGCTGCAATTAAGAACAT 1929
8617 TATATATATATACCAATGATATTATAGTACAGATGGAGCTGCAATTAAGAACAT 8558
DB TATATATATATACCAATGATATTATAGTACAGATGGAGCTGCAATTAAGAACAT 8558
1930 AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1978
DB AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1978
8557 AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8498
DB AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8498
1979 AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2038
DB AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2038
8497 AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8438
DB AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8438
2039 AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2098
DB AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2098
8437 AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8378
DB AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8378

2099 AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2158
DB AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2158
8377 AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8318
DB AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8318
2159 AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2218
DB AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2218
8317 AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8258
DB AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8258
2219 AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2278
DB AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2278
8257 AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8198
DB AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8198
2279 AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2338
DB AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2338
8197 AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8138
DB AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8138
2339 AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2391
DB AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2391
8137 AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8078
DB AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8078
2392 AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2451
DB AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2451
8077 AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8018
DB AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8018
2452 AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2511
DB AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2511
8017 AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7958
DB AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7958
2512 AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7954
DB AATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7954

RESULT 36
US-10-204-708-22/c
; Sequence 22, Application US/10204708
; Patient No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OUSK, Alexander
; APPLICANT: PIEBENROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 22
; LENGTH: 11049
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chemically treated genomic DNA (Homo sapiens)
US-10-204-708-22

Query Match 1.8%; Score 86.2; DB 4; Length 11049;
Best Local Similarity 45.9%; Pred. No. 2.9e-07;
Matches 445; Conservative 0; Mismatches 513; Indels 11; Gaps 4;
1530 TGAATATATATACCAATGATATTATAGTACAGATGGAGCTGCAATTAAGAACAT 1589
DB TGAATATATATACCAATGATATTATAGTACAGATGGAGCTGCAATTAAGAACAT 1589
10868 TGAATATATATACCAATGATATTATAGTACAGATGGAGCTGCAATTAAGAACAT 10809
DB TGAATATATATACCAATGATATTATAGTACAGATGGAGCTGCAATTAAGAACAT 10809

QY 1560 AAGCATATCAACATATATCTTGTATGTCATGCTTGTGTAATTTGACACATATATA 1649
 DB 10808 TAAATATATTTCTATTTTCTTAAATTTTAAATTCGGTTTATCGTCATACCTAA 10749
 QY 1650 AATTAACGGTCAAAATATAGATATCAAGCTTCAACATATATATAAAGAGTCAATA 1709
 DB 10748 CCTACAAAAAATATATATATATATATCA-TTTTCTAAAAAATATTTTCTT 10690
 QY 1710 CCCATGTTAATCACTCAAGTATGTAATTAATTAATTAATTAATTAATTAATTA 1769
 DB 10689 AAAAAAACAACCAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 10633
 QY 1770 AAAAAATTCCTAGCTATATATATTAACAAATTCATCCATGCTCAATTAACCAATGT 1829
 DB 10632 AACTTTAAAAAATTTCTATCAATTAATTAATTTTCTATCTTTTATTAATTAAT 10573
 QY 1830 TATCATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1889
 DB 10572 TTTTAACTTTTATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 10513
 QY 1890 CACTTTTCACTACATCATATATATATATATATATATATATATATATATATATAT 1949
 DB 10512 CAATTCATCTCAATCAATCAATTAATTAATTAATTAATTAATTAATTAATTAAT 10453
 QY 1950 AGAACCAATCAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2009
 DB 10452 -----CAATTAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 10398
 QY 2010 GTTCGCACTAATTAATTTCTATATCTGTCATATTTCAACATTAATTAATTAAT 2069
 DB 10397 AATTAATTTCTTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 10338
 QY 2070 GTTTTATATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATG 2129
 DB 10337 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 10278
 QY 2130 AATTCCTGCAATATATATATATATATATATATATATATATATATATATATAT 2189
 DB 10277 AAAACCAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 10218
 QY 2190 AATTAACCACTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2249
 DB 10217 AACTTAATAAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 10160
 QY 2250 CCTACGATGATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2309
 DB 10159 TCTTAATCTATTTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 10100
 QY 2310 TTTCAACTTCAACAACAACTCAAGTCTATATATTAATTAATTAATTAATTAAT 2369
 DB 10099 AATTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 10040
 QY 2370 CTTCACCAACATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2429
 DB 10039 CTTCCTTTAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 9980
 QY 2430 TTCACAATTCATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2489
 DB 9979 ATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 9920
 QY 2490 CAAGAAATTT 2498
 DB 9919 TATATATAT 9911

RESULT 37
 US-10-204-708-26/c
 ; Sequence 26, Application US/10204708
 ; Patent No. 6677731
 ; GENERAL INFORMATION:
 ; APPLICANT: OLEK, Alexander
 ; APPLICANT: PIEPENBROCK, Christian
 ; APPLICANT: BERLIN, Kurt

; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
 ; FILE REFERENCE: 5013.1012
 ; CURRENT APPLICATION NUMBER: US/10/204,708
 ; PRIOR FILING DATE: 2003-05-06
 ; PRIOR APPLICATION NUMBER: PCT/EP01/03971
 ; PRIOR FILING DATE: 2001-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019056.8
 ; PRIOR FILING DATE: 2000-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019173.8
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: DE 10032529.7
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: DE 10043826.1
 ; PRIOR FILING DATE: 2000-09-01
 ; NUMBER OF SEQ ID NOS: 98
 ; SEQ ID NO 26
 ; LENGTH: 6583
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
 US-10-204-708-26
 Query Match 1.8%; Score 86; DB 4; Length 6583;
 Best Local Similarity 45.0%; Pred. No. 2,9e-07;
 Matches 490; Conservative 0; Mismatches 585; Indels 13; Gaps 4;

QY 1417 ACTGAATAGAGAAAGGTAGAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1476
 DB 6526 ATTAATTAATCCAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6467
 QY 1477 TGAAGAACAGTACTGTAAACCACTGATATCAATTAATTAATTAATTAATTAAT 1536
 DB 6466 CTCACATAAATAAACCCTATATATTAATTAATTAATTAATTAATTAATTAATTA 6407
 QY 1537 ACCAATTTGATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1596
 DB 6406 ATCTATTAATAAATCACTTCTAATAAATAAATAAATAAATAAATAAATAAATAA 6347
 QY 1597 ATCAACATATATCTTATGCTATCAATCTTGTATTAATTAATTAATTAATTAAT 1656
 DB 6346 ATATACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6287
 QY 1657 GGTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1716
 DB 6286 AATCAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6227
 QY 1717 TTAACATCAACTATATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1776
 DB 6226 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6167
 QY 1777 CTAGCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1836
 DB 6166 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6115
 QY 1837 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1896
 DB 6114 CTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6055
 QY 1897 TCACTACATGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1956
 DB 6054 CAATTAAC-TAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5996
 QY 1957 ATACATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2016
 DB 5995 ACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5936
 QY 2017 CATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2076
 DB 5935 TAGGAATATCTTCTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 5879
 QY 2077 TATGCCATATGTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2136

| | | | |
|----|------|-----------------------------------------------------------------|------|
| Db | 5878 | AATTACTAAATTTTACTCTTCATATATAATATTTTACATATATAATATAATATA | 5819 |
| Oy | 2137 | TGCACAANTATGAGAGAATACCATACCGTCCAATTAATTAACAAATTTCTAATATPA | 2196 |
| Db | 5818 | TAATACATATTTACTATTTATATATATATATATATATATATATATATATATATAT | 5758 |
| Oy | 2197 | CCACTACTCTAAANATNGAACCAATAGTATATCATACCTATAGTAAGTATCCTCATAG | 2255 |
| Db | 5758 | TTTTTTTTTCAAAAAATTTTGAGATACAAAACAAAAAATATAAAAAATATCTAAAAA | 5699 |
| Oy | 2257 | CATGTATCAAATTAATPAAAGAAAAATCGTTATGTATATTAATACATTAACAAATTTCCA | 2318 |
| Db | 5688 | TTATTTTAAAAATPAAAAAAAAATTAATATTAAGAAAAACTCACTTCTCATATTTGCCAT | 5633 |
| Oy | 2317 | CTTCAACAACAAACTCAAGTCTATATATPACAAATTCAGAAATTAATATACTTCAAC | 2376 |
| Db | 5638 | ACTACAAAACCCATCTCATATATTTACTATATGATTTTCACTTAATATTTAACTTAAA | 5579 |
| Oy | 2377 | AACACATATGTATCAAATTTCTATAATACATCATATCTTTTAAAAAAGTAGCTTCAAA | 2433 |
| Db | 5578 | AACCAATTTTCATATTACACATATATCACCTTACCATAAATAAATATACCGTAAAAATACA | 5519 |
| Oy | 2437 | TTCATACATAATATTT-ATACATCGCATGGTTATTCAGAAACATTAACAAATCCAAAGA | 2499 |
| Db | 5518 | TAAACACAAAATCATTCGATTPAAAATATTTTAAAAATATAAAAAATTAATTAATCAACATCA | 5455 |
| Oy | 2496 | TTTTCAAT 2503 | |
| Db | 5458 | AAATCAAT 5451 | |

```

RESULT 38
Sequence 70, Application US/10204708
Patent No. 6677731
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENBROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
TITLE OF INVENTION: by Assessing DNA Methylation
FILE REFERENCE: 5013.1012
CURRENT APPLICATION NUMBER: US/10/204,708
CURRENT FILING DATE: 2003-05-06
PRIORITY APPLICATION NUMBER: PCT/EP01/03971
PRIORITY FILING DATE: 2001-04-06
PRIORITY APPLICATION NUMBER: DE 10019058.8
PRIORITY FILING DATE: 2000-04-06
PRIORITY APPLICATION NUMBER: DE 10019173.8
PRIORITY FILING DATE: 2000-04-07
PRIORITY APPLICATION NUMBER: DE 10032529.7
PRIORITY FILING DATE: 2000-06-30
PRIORITY APPLICATION NUMBER: DE 10043826.1
PRIORITY FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 98
SEQ ID NO 70
LENGTH: 6040
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically created genomic DNA (Homo sapiens)
US-10-204-708-70

```

| | Query Match | Similarity | Score | DB | length |
|------------|-------------|------------------------------------------------------------|-------|------------|--------|
| Best local | 629 | 44.7% | 84.7 | 07 | 6040 |
| Matches | 629 | Conservative | 0 | Mismatches | 757 |
| | | | | Indels | 20 |
| | | | | Gaps | 77 |
| Cy | 1107 | AACTTGCGATCATGTATGAAACAATAAGTAAGATGCTTAAATAATATATTCATC | | | 1166 |
| | | | | | |
| Db | 3193 | AAACGATTAACCTACAGCCCTATATATCCACACTTAAAAATATATAAAAAAATTCGC | | | 3133 |
| | | | | | |
| Oy | 1167 | AAGATATTATAGGTGTACAGATTGAGCTGCATTAAAGACATTAATGCATTTGCATGTG | | | 1228 |

[illegible]

QY 2231 CATACCTAGTAAGATATCCCTACAGCATGTATCAATAATAAAGAAAAATCGTTATGTA 2290
DB 2235 TATTTATTTCTTTAAATATTTCTCCCTAAATATCCAAAAATATAATTAATTT 2236
QY 2231 TATTTATACATTAACAAATTTCCACTTCAACACAAACAACTCAAGTCTATTATAGA 2350
DB 2235 CTATTATACATCTCTATATATATATATATTAACCAATTAACCTATATATATATATAT 2176
QY 2351 AATTAAGATATATATATATATATATATATATATATATATATATATATATATATAT 2410
DB 2175 CATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2116
QY 2411 TCCCTTTTAAAGATGACTTCAATTCATTCATTAATTAATATA 2456
DB 2115 CCATTTATATATATATATATATATATATATATATATATATATATATATATATATAT 2070

RESULT 40
US-10-204-708-29/c
Sequence 29, Application US/10204708
Patent No. 6677731
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIERENBROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
FILE REFERENCE: 5013.1012
CURRENT APPLICATION NUMBER: US/10/204.708
PRIOR FILING DATE: 2003-05-06
PRIOR APPLICATION NUMBER: PCT/EP01/03971
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: DE 10019058.8
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: DE 10019173.8
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 98
SEQ ID NO 29
LENGTH: 5666
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-29

Query Match 1.8%; Score 84.6; DB 4; Length 5666;
Best Local Similarity 44.6%; Pred. No. 5.1e-07;
Matches 1051; Conservative 0; Mismatches 1254; Indels 54; Gaps 16;

QY 149 CTCGCATTTGATATAGAGCAACCATTTTCCAAATCCCAATTTTAACATGCAACA 208
DB 2433 CTCGAAAAAATATATATATATATATATATATATATATATATATATATATATATAT 2374
QY 209 GATTAATTAATGTACAGAACACCTGTACTTATGACAGAACTATATTAATCTTATCAA 268
DB 2373 TATTCCTCAAAATATCTTAAATATACCAAAACCTATTCATAAAATCTATTAATCAAAA 2314
QY 269 GGTCTCTTACATTAACAACTGATTCAGAAATACCTCTCTCTTGTTCATTCATA 328
DB 2313 ATTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2260
QY 339 TTTTCCA-ATTATTAACATATCATAGATCTTTATTCAGATATATTAATTAATCAAT 387
DB 2259 TATTCACAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2200
QY 388 CAAAACAAAATATATATATATATATATATATATATATATATATATATATATATATAT 447
DB 2199 ACAAACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2140

QY 448 ACTCAAAATATATATAGACACTTAACCTTATATATATATATATATATATATATATATAT 507
DB 2139 TAAATATATACACCTCCACACTTTTATTTTATATATATATATATATATATATATATAT 2080
QY 508 TCAATTTTGCATTAATATACATCTGACATGATTAACATATATATATATATATATATAT 567
DB 2079 TACATTCATTTATATATATATATATATATATATATATATATATATATATATATAT 2020
QY 568 TAGATATATATATATATATATATATATATATATATATATATATATATATATATATAT 627
DB 2019 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1960
QY 628 AAAAATGAATACAGATTTTGTGACAGATCAATTTGGCAGACATATGCAATGGCACT 687
DB 1959 CCTACATACACATTTCTTTAAATCTCAATTAATTTTAAATATATATATATATATATAT 1900
QY 688 GTACAGACATGGTATATATATATATATATATATATATATATATATATATATATATAT 746
DB 1899 ACCAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1840
QY 747 TAAACACTTACCGAAAAATTTAAATATTTTACCAATATCATCTACGACACATAT 806
DB 1839 CTATCAGATATATATATATATATATATATATATATATATATATATATATATATATAT 1780
QY 807 ATGCACACGCAATATATATATATATATATATATATATATATATATATATATATATAT 862
DB 1779 TTTCTTATTTTATATATATATATATATATATATATATATATATATATATATATATAT 1720
QY 863 ACTATTTCTACACACATATATATATATATATATATATATATATATATATATATATAT 921
DB 1719 ACCACCTTTTAAATACCTTACTTAAATATTTTATTTTATATATATATATATATATAT 1660
QY 922 GTGTAAATATCTCCAAATTTATTTTGTATTAAGTATGCAATTAATCTACTCAAT 981
DB 1659 TTTACCAATTAATATATATATATATATATATATATATATATATATATATATATAT 1600
QY 982 TCCCACTGAATATGACAGAAATTTTACATATATATATATATATATATATATATATAT 1034
DB 1599 ATTTACAAACAAACAAACCAAAATCTCCATTAATATATATATATATATATATATAT 1540
QY 1035 CTGTGCAAGTTTGTATTAAGAAATTTGCAACATCTTCTGTATATATATATATATAT 1094
DB 1539 TATCTCAAAATATATATATATATATATATATATATATATATATATATATATATAT 1480
QY 1095 GTGATATATATATATATATATATATATATATATATATATATATATATATATATATAT 1154
DB 1479 TAAACACCACTACATATATATATATATATATATATATATATATATATATATATATAT 1420
QY 1155 TAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 1213
DB 1419 TAAACAAAAAATATATATATATATATATATATATATATATATATATATATATATAT 1360
QY 1214 GCATATGATATGGAATATATATATATATATATATATATATATATATATATATATAT 1273
DB 1359 TACATATACATATATATATATATATATATATATATATATATATATATATATATAT 1300
QY 1274 ATATCTAAAGCATGATGAAATTTAAACATATATATATATATATATATATATATATAT 1333
DB 1299 AATACCAATATATATATATATATATATATATATATATATATATATATATATATAT 1240
QY 1334 CTTCATATATATATATATATATATATATATATATATATATATATATATATATATATAT 1380
DB 1239 TATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 1180
QY 1381 ATGTACACAGACATATATATATATATATATATATATATATATATATATATATATAT 1439
DB 1179 ATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 1120
QY 1440 ATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 1499
DB 1119 TACATATATATATATATATATATATATATATATATATATATATATATATATATATAT 1060

QY 1500 CACTGATATCATTAAGCAAAAAAATTAAATGAGAACTACCAATTGCAATATCTATTACA 1559
DB 1059 GCACCTTCCCTGGCCGCACTACCTCCA-ACCTMAATACAAAAGAACTCCATCTCAA 1001
QY 1560 ATTATATACATGAGAAATATATACATGATTAAGCCATATCAACATATATCTTAGCTA 1619
DB 1000 AAAACCAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 943
QY 1620 TCATCCTTGTATATTTCAACCAATGATTAATAACGGTCAAAAATATAGATACAAACG 1679
DB 942 CTATATACATATTTATTAATAAACTCACTTCACTTACACACATTAATAAATCAAAATATCG 883
QY 1680 TTCACCACTATATATAAAGAGCTCAATACCCATGTTAATCCAACTAAGTAGTGA 1739
DB 882 TATACCATTAATATATTTCTATCAATACAAACCTCATATACAACTAATCTTATTA 823
QY 1740 CA-TATTACATACAAACTTAAGTATTAATTCTAGCTAATATATATTAACAA 1798
DB 822 AACGACATACATATATTTAAATTAATTAATTTCAATATTTATATATTAATTAACATTA 763
QY 1799 TTCTATCTCTCTATCTCAATTAACCAATGTTATGAGATTAATTAAGTCAATTAAC 1858
DB 762 TACTTACCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 703
QY 1859 TTTTACTTATTAATAATAATATATGTTGCACTTTTCACTACATCATGATATATATG 1918
DB 702 TATACCTTAATAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 643
QY 1919 AACATATATTTTACATTAATAATAATAATAATAATAATAATAATAATAATAATAATA 1975
DB 642 TCTAAATTTTAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 583
QY 1976 AAAACACCTATTTACATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATC 2035
DB 582 TATATATATATTTCTCAAAATATATCTCAATTAATTAATTAATTAATTAATTAATTA 523
QY 2036 CTTCACATATTTCAACACATTA-AAATATGATTAATGTTTATATGCCATTAAGTTCTTA 2094
DB 522 ATTTATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 463
QY 2095 ATTAATATGCTAA-----TTACCCAAAAGTTAAATACACATTCCTGCAACA 2144
DB 462 AAACCACTTCAATATCAATATTTCTAATTTATCACTTCAAAATCTATTAATAATAACA 403
QY 2145 TATGAGAAAGATACCAATACAGTCAATTAATTAATTAATTAATTAATTAATTAATC 2204
DB 402 ACTAAACCCAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAA 343
QY 2205 TCTAAACATAGACCAATAGATATCTACTAAGTAGATATCCCTAGCAAGATGATC 2264
DB 342 ATTTAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 283
QY 2265 AAATAATAAAGAAATCGTTATGATATTTATTAATTAATTAATTAATTAATTAATTA 2324
DB 282 CAATAAATCAACAATCTCTAACAATAATAATAATAATAATAATAATAATAATAATA 223
QY 2325 ACAAAACCTCAA-AGTCTATATTAACAATTAAGATTAATTAATTAATTAATTAATTA 2383
DB 222 ACCAATATCAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 163
QY 2384 AGTTATCAATCTATATATCAATCAATATCTTTTAAAGATGATGATCAATTCATATC 2443
DB 162 ATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 103
QY 2444 ATTAATATTTATTAATGCA 2462
DB 102 ATTAATTTCTCAAAAATCA 84

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2004, 18:26:43 ; Search time 1530 Seconds
(without alignments)

11633.871 Million cell updates/sec

Title: US-10-081-051-52
Perfect score: 4750
Sequence: 1 gatccgtactatctagg.....ccacattgactatgac 4750

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2470430 seqs, 1873665578 residues
Total number of hits satisfying chosen parameters: 4940860

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|---------|----------------------|-------------------|
| 1 | 4750 | 100.0 | 4750 | US-10-081-051-52 | Sequence 52, Appl |
| 2 | 1416 | 29.8 | 1416 | US-10-081-051-56 | Sequence 56, Appl |
| 3 | 984 | 20.7 | 984 | US-10-081-051-54 | Sequence 54, Appl |
| 4 | 587 | 12.6 | 587 | US-10-081-051-57 | Sequence 57, Appl |
| 5 | 371 | 7.8 | 371 | US-10-081-051-53 | Sequence 53, Appl |
| 6 | 269.2 | 5.7 | 1371 | US-10-010-160-5 | Sequence 5, Appl |
| 7 | 207.8 | 4.4 | 3673778 | US-10-312-841-1 | Sequence 1, Appl |
| 8 | 187.6 | 3.9 | 1017 | US-10-282-122A-16370 | Sequence 16370, A |
| 9 | 175.2 | 3.7 | 1029 | US-10-282-122A-17353 | Sequence 17353, A |
| 10 | 170.6 | 3.6 | 996 | US-10-282-122A-35500 | Sequence 35500, A |
| 11 | 168.6 | 3.5 | 989 | US-09-815-242-8974 | Sequence 8974, Ap |
| 12 | 168.4 | 3.5 | 999 | US-10-282-122A-15395 | Sequence 15395, A |
| 13 | 165.4 | 3.5 | 999 | US-10-282-122A-34158 | Sequence 34158, A |
| 14 | 155.2 | 3.3 | 7814 | US-10-221-714A-252 | Sequence 252, App |
| 15 | 155 | 3.3 | 1065 | US-10-282-122A-21770 | Sequence 21770, A |

| | | | | | | |
|------|-------|-----|---------|----|----------------------|-------------------|
| C 16 | 155 | 3.3 | 5641 | 14 | US-10-311-455-1369 | Sequence 1369, Ap |
| C 17 | 154.2 | 3.2 | 918 | 8 | US-08-781-986A-513 | Sequence 513, App |
| C 18 | 154.2 | 3.2 | 918 | 12 | US-10-329-624-513 | Sequence 513, App |
| C 19 | 153 | 3.2 | 15548 | 14 | US-10-311-455-2128 | Sequence 2128, Ap |
| C 20 | 148.8 | 3.1 | 6314 | 14 | US-10-240-452-14 | Sequence 14, Appl |
| C 21 | 148.4 | 3.1 | 1020 | 12 | US-10-282-122A-6338 | Sequence 6338, Ap |
| C 22 | 148.4 | 3.1 | 1023 | 9 | US-09-815-242-6755 | Sequence 3863, Ap |
| C 23 | 147.2 | 3.1 | 1026 | 9 | US-09-815-242-3869 | Sequence 3869, Ap |
| C 24 | 147.2 | 3.1 | 6132 | 14 | US-10-311-455-836 | Sequence 836, App |
| C 25 | 146.2 | 3.1 | 319630 | 15 | US-10-398-221-7 | Sequence 7, Appl |
| C 26 | 145.2 | 3.1 | 3013208 | 15 | US-10-398-221-2058 | Sequence 2058, Ap |
| C 27 | 145.2 | 3.1 | 11745 | 14 | US-10-240-453-206 | Sequence 206, App |
| C 28 | 144.2 | 3.0 | 7597 | 14 | US-10-311-455-886 | Sequence 986, App |
| C 29 | 141.6 | 3.0 | 3673778 | 14 | US-10-312-841-2 | Sequence 2, Appl |
| C 30 | 139.8 | 2.9 | 2791 | 15 | US-10-398-221-3618 | Sequence 3618, Ap |
| C 31 | 139 | 2.9 | 6175 | 14 | US-10-311-455-1280 | Sequence 1280, Ap |
| C 32 | 137 | 2.9 | 18154 | 14 | US-10-311-455-227 | Sequence 227, App |
| C 33 | 136.8 | 2.9 | 1008 | 9 | US-09-815-242-6968 | Sequence 6968, Ap |
| C 34 | 136.8 | 2.9 | 1008 | 12 | US-10-282-122A-22052 | Sequence 22052, A |
| C 35 | 136.8 | 2.9 | 1830121 | 14 | US-10-329-660-1 | Sequence 1, Appl |
| C 36 | 136.8 | 2.9 | 1830121 | 15 | US-10-329-670-1 | Sequence 1, Appl |
| C 37 | 136.2 | 2.9 | 11691 | 14 | US-10-311-455-2213 | Sequence 2213, Ap |
| C 38 | 136 | 2.9 | 17594 | 14 | US-10-311-455-1999 | Sequence 1999, Ap |
| C 39 | 135.8 | 2.9 | 6831 | 14 | US-10-311-455-1460 | Sequence 1460, Ap |
| C 40 | 135.4 | 2.9 | 11691 | 14 | US-10-311-455-2214 | Sequence 2214, Ap |
| C 41 | 134.8 | 2.8 | 15387 | 14 | US-10-311-455-157 | Sequence 157, App |
| C 42 | 134.4 | 2.8 | 6106 | 12 | US-10-221-714A-151 | Sequence 151, App |
| C 43 | 134.4 | 2.8 | 6106 | 14 | US-10-311-455-1445 | Sequence 1445, Ap |
| C 44 | 134.4 | 2.8 | 6106 | 16 | US-10-257-166-113 | Sequence 113, App |
| C 45 | 134 | 2.8 | 7047 | 14 | US-10-240-453-260 | Sequence 260, App |

ALIGNMENTS

RESULT 1
US-10-081-051-52
Sequence 52, Application US/10081051
Publication No. US2003004422A1
GENERAL INFORMATION:
APPLICANT: Barbet, Anthony F.
APPLICANT: Whitmore, William M.
APPLICANT: Kamper, Sondra M.
APPLICANT: Simbi, Bigboy H.
APPLICANT: Ganta, Roman R.
APPLICANT: Moreland, Annie L.
APPLICANT: Mwangi, Duncan M.
APPLICANT: McGuire, Travis C.
APPLICANT: Mahan, Suman M.
TITLE OF INVENTION: Ehrlichia Ruminantium Polypeptides, Antigens, Polynucleotides, a
FILE REFERENCE: US-295XCI
CURRENT APPLICATION NUMBER: US/10/081,051
CURRENT FILING DATE: 2002-02-20
PRIORITY APPLICATION NUMBER: US 60/269,944
PRIOR FILING DATE: 2001-02-20
NUMBER OF SEQ ID NOS: 117
SOFTWARE: PatentIn version 3.1
SEQ ID NO 52
LENGTH: 4750
TYPE: DNA
ORGANISM: Ehrlichia ruminantium (formerly Cowdria ruminantium)
US-10-081-051-52

Query Match 100.0%; Score 4750; DB 14; Length 4750;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4750; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATCCGCTAATCTTATGATGCTTTTAAAGTGGCTGATATAAATAAACT 60
DB 1 GATCCGCTAATCTTATGATGCTTTTAAAGTGGCTGATATAAATAAACT 60
QY 61 GCTTGACTGCTATATATAGTACATACGCTGATAGTAAAGAAAAATTCTCAAGCAGAC 120

```
Db 61 GCTTGACTGTCTATATAGTACATACGTGATGATGAGAAAAATTTCTCAAGAGAC 120
Qy 121 AGTACCTTGCTGATGATGATGCTGCTTCACTCCACATTTGATGAGCAAAACATTTT 180
Db 121 AGTACCTTGCTGATGATGATGCTGCTTCACTCCACATTTGATGAGCAAAACATTTT 180
Qy 181 CTCAATCTGCAATTTTAAAGTACGACAAAGATTTATGTAACAGAACCCCTGATC 240
Db 181 CTCAATCTGCAATTTTAAAGTACGACAAAGATTTATGTAACAGAACCCCTGATC 240
Qy 241 TTATGCAAGAACTCTATPAACTTATGCAAGGCTCTTACATPAGAAACATGTAATCAA 300
Db 241 TTATGCAAGAACTCTATPAACTTATGCAAGGCTCTTACATPAGAAACATGTAATCAA 300
Qy 301 GAATACCTCTCTCTGTCATGATGATTTTCCATTTTCCATTTTCCATTTTCCATTTTCC 360
Db 301 GAATACCTCTCTCTGTCATGATGATTTTCCATTTTCCATTTTCCATTTTCCATTTTCC 360
Qy 361 TTATGCAATPACATTTTAAATTCATCAAAACAAATATATAGTATCTATTTATATA 420
Db 361 TTATGCAATPACATTTTAAATTCATCAAAACAAATATATAGTATCTATTTATATA 420
Qy 421 AAAATCTCTGATGCTTTAGCAAGACCTACAAATTTATGACACCTTAATTTAA 480
Db 421 AAAATCTCTGATGCTTTAGCAAGACCTACAAATTTATGACACCTTAATTTAA 480
Qy 481 TATATTTCAATTTATTTATCTACTCATATTTTATGATATATGATCTGACATGAT 540
Db 481 TATATTTCAATTTATTTATCTACTCATATTTTATGATATATGATCTGACATGAT 540
Qy 541 TAAATATGATPACATTTTCTTAACTATGATATATATATATATATATATTTT 600
Db 541 TAAATATGATPACATTTTCTTAACTATGATATATATATATATATATATTTT 600
Qy 601 GACTGTTCTGATATATATATATATATATATATATATATATATATATATATAT 660
Db 601 GACTGTTCTGATATATATATATATATATATATATATATATATATATATATAT 660
Qy 661 ATTGTCAGACATATGATATGATGATGATGATGATGATGATGATGATGATGATG 720
Db 661 ATTGTCAGACATATGATATGATGATGATGATGATGATGATGATGATGATGATG 720
Qy 721 ACGTATCATAGAAATTTACATATATATATATATATATATATATATATATATAT 780
Db 721 ACGTATCATAGAAATTTACATATATATATATATATATATATATATATATATAT 780
Qy 781 CACATATCATATACAGACAAATATATATATATATATATATATATATATATATAT 840
Db 781 CACATATCATATACAGACAAATATATATATATATATATATATATATATATATAT 840
Qy 841 CAACAATACATGATATTTCTTAACTATTTCTTCAACAACAATATGACACATATG 900
Db 841 CAACAATACATGATATTTCTTAACTATTTCTTCAACAACAATATGACACATATG 900
Qy 901 AATPACACACAAACAGATATGATGATATATATATATATATATATATATATATAT 960
Db 901 AATPACACACAAACAGATATGATGATATATATATATATATATATATATATATAT 960
Qy 961 CGAAATATACATCTCAAAATTTTCCAGTAAATGACGAAAGAAATTTTCAATATAT 1020
Db 961 CGAAATATACATCTCAAAATTTTCCAGTAAATGACGAAAGAAATTTTCAATATAT 1020
Qy 1021 AATTTTATATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
Db 1021 AATTTTATATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
Qy 1081 TATGATCTGCTGATGATATATATATATATATATATATATATATATATATATAT 1140
Db 1081 TATGATCTGCTGATGATATATATATATATATATATATATATATATATATATATAT 1140
Qy 1141 TGATGTTCTAAATATATATATATATATATATATATATATATATATATATATAT 1200
Db 1141 TGATGTTCTAAATATATATATATATATATATATATATATATATATATATATAT 1200

Db 1141 TGATGTTCTAAATATATATATATATATATATATATATATATATATATATATAT 1200
Qy 1201 AAGAACATATATGCAATTTGATGATGATGATGATGATGATGATGATGATGATG 1260
Db 1201 AAGAACATATATGCAATTTGATGATGATGATGATGATGATGATGATGATGATG 1260
Qy 1261 TGTGCTGCTGATATATATATATATATATATATATATATATATATATATATAT 1320
Db 1261 TGTGCTGCTGATATATATATATATATATATATATATATATATATATATATAT 1320
Qy 1321 TCATTCATATGATCTTCAATATATATATATATATATATATATATATATATATAT 1380
Db 1321 TCATTCATATGATCTTCAATATATATATATATATATATATATATATATATATAT 1380
Qy 1381 ATGTCACACAGACATTTGACCAATATATATATATATATATATATATATATATAT 1440
Db 1381 ATGTCACACAGACATTTGACCAATATATATATATATATATATATATATATATAT 1440
Qy 1441 TATATATATATATATATATATATATATATATATATATATATATATATATATAT 1500
Db 1441 TATATATATATATATATATATATATATATATATATATATATATATATATATAT 1500
Qy 1501 ACTGATATATATATATATATATATATATATATATATATATATATATATATAT 1560
Db 1501 ACTGATATATATATATATATATATATATATATATATATATATATATATATAT 1560
Qy 1561 TTTATATATATATATATATATATATATATATATATATATATATATATATATAT 1620
Db 1561 TTTATATATATATATATATATATATATATATATATATATATATATATATATAT 1620
Qy 1621 CATCTTGTATATATATATATATATATATATATATATATATATATATATATAT 1680
Db 1621 CATCTTGTATATATATATATATATATATATATATATATATATATATATATAT 1680
Qy 1681 TCACACATATATATATATATATATATATATATATATATATATATATATATAT 1740
Db 1681 TCACACATATATATATATATATATATATATATATATATATATATATATATAT 1740
Qy 1741 ATATATATATATATATATATATATATATATATATATATATATATATATATAT 1800
Db 1741 ATATATATATATATATATATATATATATATATATATATATATATATATATAT 1800
Qy 1801 CCTATCTGCTATCTCAATTTTACATATATATATATATATATATATATATATAT 1860
Db 1801 CCTATCTGCTATCTCAATTTTACATATATATATATATATATATATATATATAT 1860
Qy 1861 TTAATATATATATATATATATATATATATATATATATATATATATATATATAT 1920
Db 1861 TTAATATATATATATATATATATATATATATATATATATATATATATATATAT 1920
Qy 1921 CATATATATATATATATATATATATATATATATATATATATATATATATATAT 1980
Db 1921 CATATATATATATATATATATATATATATATATATATATATATATATATATAT 1980
Qy 1981 ACCTATATATATATATATATATATATATATATATATATATATATATATATAT 2040
Db 1981 ACCTATATATATATATATATATATATATATATATATATATATATATATATAT 2040
Qy 2041 CATATATATATATATATATATATATATATATATATATATATATATATATATAT 2100
Db 2041 CATATATATATATATATATATATATATATATATATATATATATATATATATAT 2100
Qy 2101 TTGCTATATATATATATATATATATATATATATATATATATATATATATATAT 2160
Db 2101 TTGCTATATATATATATATATATATATATATATATATATATATATATATATAT 2160
Qy 2161 TACAGCTCAATATATATATATATATATATATATATATATATATATATATATAT 2220
Db 2161 TACAGCTCAATATATATATATATATATATATATATATATATATATATATATAT 2220
Qy 2221 AATATATATATATATATATATATATATATATATATATATATATATATATATAT 2280
Db 2221 AATATATATATATATATATATATATATATATATATATATATATATATATATAT 2280
```


[illegible]

| | | | |
|----|------|-----------------------------------------------------------------|------|
| QY | 3361 | CGAGTAGTACTGCTGTTATTAATCAATCACTCTCAAAATTAGCTAAACAGGGAGTTAC | 3420 |
| Db | 3361 | CGAGTAGTACTGCTGTTATTAATCAATCACTCTCAAAATTAGCTAAACAGGGAGTTAC | 3420 |
| QY | 3421 | GACTATTTCAATATCTCAAAATGCTGATTTCAACTCATATTTCTCTAGTAATCTAAC | 3480 |
| Db | 3421 | GACTATTTCAATATCTCAAAATGCTGATTTCAACTCATATTTCTCTAGTAATCTAAC | 3480 |
| QY | 3481 | AGACTCTATGACTCTCTTTACAACTAACCTTAATCTCTCTGTGAAAAGGCTTTTCAT | 3540 |
| Db | 3481 | AGACTCTATGACTCTCTTTACAACTAACCTTAATCTCTCTGTGAAAAGGCTTTTCAT | 3540 |
| QY | 3541 | ATTAATCATAGAACCCCAATGCAAGAGCTTTACGACAGTGGCAATATTAATCCATGCCACT | 3600 |
| Db | 3541 | ATTAATCATAGAACCCCAATGCAAGAGCTTTACGACAGTGGCAATATTAATCCATGCCACT | 3600 |
| QY | 3601 | AATCATTAATAAGGCAAAATAGAGATGACCTTTCTTTAAGCTTTCCAGTACACTTTAATCC | 3660 |
| Db | 3601 | AATCATTAATAAGGCAAAATAGAGATGACCTTTCTTTAAGCTTTCCAGTACACTTTAATCC | 3660 |
| QY | 3661 | ATCAATATCAGATCCTCTTAACAATATATCCAAATATACAACTAGGCTCTTTTTCATA | 3720 |
| Db | 3661 | ATCAATATCAGATCCTCTTAACAATATATCCAAATATACAACTAGGCTCTTTTTCATA | 3720 |
| QY | 3721 | AGCATCTGTGATCGCGGATTAACCATCAACGCTGAATTTAGTGACATTAATATCATCACT | 3780 |
| Db | 3721 | AGCATCTGTGATCGCGGATTAACCATCAACGCTGAATTTAGTGACATTAATATCATCACT | 3780 |
| QY | 3781 | TAAATATCTTTTATTTAGATTTCGATATCACTTCAATCAACAACATAACTTTCAGA | 3840 |
| Db | 3781 | TAAATATCTTTTATTTAGATTTCGATATCACTTCAATCAACAACATAACTTTCAGA | 3840 |
| QY | 3841 | AATATACATCTTCTCTGACATTTCAAAATCCTGTGCAATAATTAACCTACATATCAAA | 3900 |
| Db | 3841 | AATATACATCTTCTCTGACATTTCAAAATCCTGTGCAATAATTAACCTACATATCAAA | 3900 |
| QY | 3901 | TGAAATATCTGAAAACAATGTACTCCTTTAAGTACACCTCAATATATCTACACACACA | 3960 |
| Db | 3901 | TGAAATATCTGAAAACAATGTACTCCTTTAAGTACACCTCAATATATCTACACACACA | 3960 |
| QY | 3961 | CTAATGATCAACAACCAATATTAATTAATGCTATTTTCTTACTATTAATTAACAATAT | 4020 |
| Db | 3961 | CTAATGATCAACAACCAATATTAATTAATGCTATTTTCTTACTATTAATTAACAATAT | 4020 |
| QY | 4021 | TATATAAATTGTTACTATATTTACTAATTTATACAGATATGTTTCTATATATCCAAA | 4080 |
| Db | 4021 | TATATAAATTGTTACTATATTTACTAATTTATACAGATATGTTTCTATATATCCAAA | 4080 |
| QY | 4081 | AAGAAAAATAAAAACCTTATATAAAAAAGTTATCATCACTTGTTAAACCTATGTAAT | 4140 |
| Db | 4081 | AAGAAAAATAAAAACCTTATATAAAAAAGTTATCATCACTTGTTAAACCTATGTAAT | 4140 |
| QY | 4141 | AATATATATTAATTAATTTTCAATATATTAACCTTCCGTGAACAGCATGCTATAGTTT | 4200 |
| Db | 4141 | AATATATATTAATTAATTTTCAATATATTAACCTTCCGTGAACAGCATGCTATAGTTT | 4200 |
| QY | 4201 | TCCATACTAAAAATCTCACTGGAAGCTTTATAGATATTCATTAATTCAGCCCAACAGAT | 4260 |
| Db | 4201 | TCCATACTAAAAATCTCACTGGAAGCTTTATAGATATTCATTAATTCAGCCCAACAGAT | 4260 |
| QY | 4261 | AAGAAATCAATTAATGATATTTTATATTTTAAACCATAGACAAATATATATCCACCAAT | 4320 |
| Db | 4261 | AAGAAATCAATTAATGATATTTTATATTTTAAACCATAGACAAATATATATCCACCAAT | 4320 |
| QY | 4321 | GACTATCAACATATGATATGATGAACCAATACCAACAAAGAGATTTATCATTTGACTA | 4380 |
| Db | 4321 | GACTATCAACATATGATATGATGAACCAATACCAACAAAGAGATTTATCATTTGACTA | 4380 |
| QY | 4381 | CTTACAGATTAATTAATCTAGATGAGTTAAAGCAACATCTTCATGATTTGAAGTCAAAATCAT | 4440 |
| Db | 4381 | CTTACAGATTAATTAATCTAGATGAGTTAAAGCAACATCTTCATGATTTGAAGTCAAAATCAT | 4440 |
| QY | 4441 | CCCGTATATTAACAATATATCAACAATAATATGTCCAAAACCTATCAATATAGAAATTTTGT | 4500 |

QY 3727 CTGTGCGGATTAACCATCATCTGCTAATTAGTGCACATTAATTCATCTTAATAT 3786
DB 1261 CTGTGCGGATTAACCATCATCTGCTAATTAGTGCACATTAATTCATCTTAATAT 1320
QY 3787 ATCTTTTATGATTTTCTGATATCACTTATCATCAACATTAATCTTCAAGAAATTA 3846
DB 1321 ATCTTTTATGATTTTCTGATATCACTTATCATCAACATTAATCTTCAAGAAATTA 1380
QY 3847 CAATCTTCTTGACATTTCAAAATCCTGTGCCAT 3882
DB 1381 CAATCTTCTTGACATTTCAAAATCCTGTGCCAT 1416

RESULT 3
US-10-081-051-54
Sequence 54, Application US/10081051
Publication No. US20030044422A1
GENERAL INFORMATION:

APPLICANT: Barbet, Anthony F.
APPLICANT: Whitmire, William M.
APPLICANT: Kamper, Sondra M.
APPLICANT: Simbi, Bigboy H.
APPLICANT: Ganta, Roman R.
APPLICANT: Moreland, Annie L.
APPLICANT: Mwangi, Duncan M.
APPLICANT: McGuire, Travis C.

APPLICANT: Mahan, Suman M.
TITLE OF INVENTION: Ehrlichia Ruminantium Polypeptides, Antigens, Polynucleotides, and
METHOD OF INVENTION: Methods of Use

FILE REFERENCE: US-299XCI

CURRENT APPLICATION NUMBER: US/10/081, 051

PRIOR FILING DATE: 2002-02-20

PRIOR APPLICATION NUMBER: US 60/269, 944

NUMBER OF SEQ ID NOS: 117

SOFTWARE: Patent version 3.1

SEQ ID NO 54

LENGTH: 984

TYPE: DNA

ORGANISM: Ehrlichia ruminantium (formerly Cowdria ruminantium)

FEATURE:

NAME/KEY: CDS

LOCATION: (1)..(984)

OTHER INFORMATION: Corresponds to SEQ ID NO:52, nucleotides 632..1615

OTHER INFORMATION: Hypothetical glycerol 3-phosphate dehydrogenase

OTHER INFORMATION: Product = "19tworf2"

US-10-081-051-54

Query Match

Best Local Similarity 100.0%; Score 984; DB 14; Length 984;

Matches 984; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 632 ATGAATACAGTATTTTGGTGCAGATCATTGGACAGAGCAATGCAATGGACCTGCA 691
DB 1 ATGAATACAGTATTTTGGTGCAGATCATTGGACAGAGCAATGCAATGGACCTGCA 60
QY 692 GCAATGATATATCACTTAATTTGGGAGAGTATCATGAAATATTAACATATAAC 751
DB 61 GCAATGATATATCACTTAATTTGGGAGAGTATCATGAAATATTAACATATAAC 120
QY 752 ACTTACGAAAAAATTTAAATATTTTACCCACATATCATCTACAGCAACATATATGCA 811
DB 121 ACTTACGAAAAAATTTAAATATTTTACCCACATATCATCTACAGCAACATATATGCA 180
QY 812 ACCGCAATATATGAGAAATTTATCTGACAAACATATCATGATTTACTTACTTCT 871
DB 181 ACCGCAATATATGAGAAATTTATCTGACAAACATATCATGATTTACTTACTTCT 240
QY 872 ACACAAATATATGAGCAATATGATATGACAAATATGACAAACAGCATATGTAATAAT 931
DB 241 ACACAAATATATGAGCAATATGATATGACAAATATGACAAACAGCATATGTAATAAT 300

QY 932 ACTCCAAATTAATTTGATGAAAGGATCGAAATTAATCATCTCAATTTCCAGTGA 991
DB 301 ACTCCAAATTAATTTGATGAAAGGATCGAAATTAATCATCTCAATTTCCAGTGA 360
QY 992 ATAGCAGAAAGAAATTTTCAATATATCAATTTTATATCTCTGGGCAAGTTTCT 1051
DB 361 ATAGCAGAAAGAAATTTTCAATATATCAATTTTATATCTCTGGGCAAGTTTCT 420
QY 1052 AAAGAAATTCAGAAATCTCTCTGTGTATGATGATCTGTGTGTATTAAGAACTT 1111
DB 421 AAAGAAATTCAGAAATCTCTCTGTGTATGATGATCTGTGTGTATTAAGAACTT 480
QY 1112 GGTGATCTATGATGAAACAAATTAAGTATGATGATGATGATGATGATGATGAT 1171
DB 481 GGTGATCTATGATGAAACAAATTAAGTATGATGATGATGATGATGATGATGAT 540
QY 1172 ATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1231
DB 541 ATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
QY 1232 ATGCTGGAAGAAATTTAGATATATGCTGTGTCTATCTTATTAAGGATGAT 1291
DB 601 ATGCTGGAAGAAATTTAGATATATGCTGTGTCTATCTTATTAAGGATGAT 660
QY 1292 GAAATTAAGAACTATATATGACAAAGAAATCATTCATGATGATCTCATATTAAT 1351
DB 661 GAAATTAAGAACTATATATGACAAAGAAATCATTCATGATGATCTCATATTAAT 720
QY 1352 CCATCATGCTTTGGAGATCTAATATTAACATGATGACAGAAACATTCAGCAAT 1411
DB 721 CCATCATGCTTTGGAGATCTAATATTAACATGATGACAGAAACATTCAGCAAT 780
QY 1412 TTGAGCTGAATATGAAAGAGTGAATATTAATTAATTAATTAATTAATTAATTA 1471
DB 781 TTGAGCTGAATATGAAAGAGTGAATATTAATTAATTAATTAATTAATTAATTA 840
QY 1472 CTGTGGAAGAAACAGTATCTGTAACCACTGATATGATTAAGCAAAATTAAT 1531
DB 841 CTGTGGAAGAAACAGTATCTGTAACCACTGATATGATTAAGCAAAATTAAT 900
QY 1532 GAATCAACATTTGATGATCTATTTTACAAATTAATTAATTAATTAATTAATTA 1591
DB 901 GAATCAACATTTGATGATCTATTTTACAAATTAATTAATTAATTAATTAATTA 960
QY 1592 GCCATATCAACATTTATCTTAG 1615
DB 961 GCCATATCAACATTTATCTTAG 984

RESULT 4

US-10-081-051-57

Sequence 57, Application US/10081051

Publication No. US20030044422A1

GENERAL INFORMATION:

APPLICANT: Barbet, Anthony F.

APPLICANT: Whitmire, William M.

APPLICANT: Kamper, Sondra M.

APPLICANT: Simbi, Bigboy H.

APPLICANT: Ganta, Roman R.

APPLICANT: Moreland, Annie L.

APPLICANT: Mwangi, Duncan M.

APPLICANT: McGuire, Travis C.

APPLICANT: Mahan, Suman M.

TITLE OF INVENTION: Ehrlichia Ruminantium Polypeptides, Antigens, Polynucleotides, and

METHOD OF INVENTION: Methods of Use

FILE REFERENCE: US-299XCI

CURRENT APPLICATION NUMBER: US/10/081, 051

PRIOR FILING DATE: 2002-02-20

PRIOR APPLICATION NUMBER: US 60/269, 944

NUMBER OF SEQ ID NOS: 117

SOFTWARE: Patent version 3.1

SEQ ID NO 57

```

; LENGTH: 597
; TYPE: DNA
; ORGANISM: Ehrlichia ruminantium (formerly Cowdria ruminantium)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(597)
; OTHER INFORMATION: Complement to SEQ ID NO:52, nucleotides 4154..>4750
; OTHER INFORMATION: Hypothetical cell division protein ftsQ
; OTHER INFORMATION: Product = "19hworf41"
US-10-081-051-57

Query Match          12.6%; Score 597; DB 14; Length 597;
Best Local Similarity 100.0%; Pred. No. 2.5e-85;
Matches 597; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4154 TTAATTCATTAATTAACCTTCCCTGGAAACAGCATGTGTATAGTTTCCATACATAAA 4213
DB 1 TTAATTCATTAATTAACCTTCCCTGGAAACAGCATGTGTATAGTTTCCATACATAAA 60
QY 4214 CTCACTTGAAAGCTTTATAGATATTCATTAATTCAGCCCAAGCATAGATCATTTAT 4273
DB 61 CTCACTTGAAAGCTTTATAGATATTCATTAATTCAGCCCAAGCATAGATCATTTAT 120
QY 4274 AGGTAATTTATATTAATTAACCATAGCAAAATATATCCCAATGACTATCAACATA 4333
DB 121 AGGTAATTTATATTTTAACCATAGCAAAATATATCCCAATGACTATCAACATA 180
QY 4334 TGTGATAGATGAACCAATACCAACCAACAAAGATATTAATGACTACTTCAAGTAAA 4393
DB 181 TGTGATAGATGAACCAATACCAACCAACAAAGATATTAATGACTACTTCAAGTAAA 240
QY 4394 ATCTAGATGATGAACCAACATCTCATGATATTAAGTCAATATCCCGATATTTACA 4453
DB 241 ATCTAGATGATGAACCAACATCTCATGATATTAAGTCAATATCCCGATATTTACA 300
QY 4454 ATTATCAACAATATATATGTCGAAACTATCATATGATAGATATTTGTTATCAATGATCA 4513
DB 301 ATTATCAACAATATATATGTCGAAACTATCATATGATAGATATTTGTTATCAATGATCA 360
QY 4514 ATTGCAAAAGCAATATCTTGAACCTGTTATCTGTAAAGTATTTGGTAAGAGCTCTT 4573
DB 361 ATTGCAAAAGCAATATCTTGAACCTGTTATCTGTAAAGTATTTGGTAAGAGCTCTT 420
QY 4574 AACTGAGGCAATTTTGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4633
DB 421 AACTGAGGCAATTTTGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 4634 AGGCACAAAAGATAGACCTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 4693
DB 481 AGGCACAAAAGATAGACCTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
QY 4694 TTCATTACCATCATCATATTTTATCAACAGTAAATCCAAATGACTAATTTGATC 4750
DB 541 TTCATTACCATCATCATATTTTATCAACAGTAAATCCAAATGACTAATTTGATC 597

RESULT 5
US-10-081-051-53
; Sequence 53, Application US/10081051
; Publication No. US20030044422A1
; GENERAL INFORMATION:
; APPLICANT: Barber, Anthony F.
; APPLICANT: Whitmore, William M.
; APPLICANT: Kamper, Sandra M.
; APPLICANT: Simbi, Bigboy H.
; APPLICANT: Ganta, Roman R.
; APPLICANT: Moreland, Annie L.
; APPLICANT: Mwambi, Duncan M.
; APPLICANT: McGuire, Travis C.
; APPLICANT: Mahan, Susan M.
; TITLE OF INVENTION: Ehrlichia Ruminantium Polypeptides, Antigens, Polynucleotides, an
; FILE REFERENCE: US-299XCI
```

```

; CURRENT APPLICATION NUMBER: US/10/081,051
; CURRENT FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: US 60/269,944
; PRIOR FILING DATE: 2001-02-20
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: Patent version 3.1
; SEQ ID NO 53
; LENGTH: 371
; TYPE: DNA
; ORGANISM: Ehrlichia ruminantium (formerly Cowdria ruminantium)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(371)
; OTHER INFORMATION: Complement to SEQ ID NO:52, nucleotides <1..371
; OTHER INFORMATION: Hypothetical integrase/recombinase
; OTHER INFORMATION: Product = "19hworf11"
US-10-081-051-53

Query Match          7.8%; Score 371; DB 14; Length 371;
Best Local Similarity 100.0%; Pred. No. 1.4e-49;
Matches 371; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATCCTGCTAATCTTTAGGTAATGTTTATTTTAAAGTGTGCTAGATAATAAACT 60
DB 1 GATCCTGCTAATCTTTAGGTAATGTTTATTTTAAAGTGTGCTAGATAATAAACT 60
QY 61 GCTTGACTGTCTATATAGTACATACGATAGATAGATAGATAGATAGATAGATAGATAGAT 120
DB 61 GCTTGACTGTCTATATAGTACATACGATAGATAGATAGATAGATAGATAGATAGATAGAT 120
QY 61 GCTTGACTGTCTATATAGTACATACGATAGATAGATAGATAGATAGATAGATAGATAGAT 180
DB 121 AGTACCTTGATGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
QY 121 AGTACCTTGATGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
DB 181 CTCAATCTCCTATTTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 181 CTCAATCTCCTATTTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
DB 241 TTATGCAAGAACTCTATTAACCTTAACAGCTCTTTCATTAAGAAACAACTGATTTCAA 300
QY 241 TTATGCAAGAACTCTATTAACCTTAACAGCTCTTTCATTAAGAAACAACTGATTTCAA 360
DB 301 GAATACCTCTCTCTTGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
QY 301 GAATACCTCTCTCTTGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
DB 361 TTATTCGACAT 371
QY 361 TTATTCGACAT 371
DB 361 TTATTCGACAT 371

RESULT 6
US-10-010-160-5/c
; Sequence 5, Application US/10010160
; Publication No. US20030103999A1
; GENERAL INFORMATION:
; APPLICANT: Rosey, Everett L.
; APPLICANT: Strugnell, Richard A.
; APPLICANT: Good, Robert T.
; APPLICANT: King, Kendall W.
; TITLE OF INVENTION: NOVEL THERAPEUTIC COMPOSITIONS FOR
; FILE REFERENCE: DAVI110.001AUS
; CURRENT APPLICATION NUMBER: US/10/010,160
; CURRENT FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: AU 281381
; PRIOR FILING DATE: 2000-11-10
; PRIOR APPLICATION NUMBER: US 60/249,596
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1371
```


| | | | | |
|----|--|---------|-----------------------------------------------------------------|----------|
| OY | | 379 | AAATGCATCAAAAACAAATAATAGTACTACTTAATTATATAAAAAACCCTAATGCT | 438 |
| Dd | | 1715369 | AATATATATATATATATATATATATATATATATATATATATATATATATATATATAT | 17152908 |
| OY | | 439 | TTAGCAGACTCACCAATATATATGAGACCTTAACCTATATATATATTTCAAATTTATTA | 498 |
| Dd | | 1715289 | ATATATATATATATATATATATATATATATATATATATATATATATATATATATAT | 17152300 |
| OY | | 499 | TACTACTCATCATATTTTGACATTAATATACATCTGACATGTATTAACAATAGTAACACTT | 558 |
| Dd | | 1715229 | AATATATATATATATATATATATATATATATATATATATATATATATATATATATAT | 17151707 |
| OY | | 559 | TCTTACACTGTATATATATACAAATATATATATATATTTTGACTGGTTTCTTGATATA | 618 |
| Dd | | 1715169 | TATATATATATATATATATATATATATATATATATATATATATATATATATATATAT | 17151100 |
| OY | | 619 | AATTAGTAAAAAATGAATAATCAGTATTTTAGTGTCAGATCCATTTGGCAGCAAAATAGC | 678 |
| Dd | | 1715109 | AAATATATATATATATATATATATATATATATATATATATATATATATATATATAT | 17150506 |
| OY | | 679 | AATTGCACTGTCAGACATCGGTATATCGTTTACTTATGGGGAGCTGATCATAGAAATAT | 738 |
| Dd | | 1715049 | TAAAAATATATATATATATATATATATATATATATATATATATATATATATATATAT | 17149917 |
| OY | | 739 | TACACATATTAACACTTACCGAAAAAATTTAAAAATTTACCCATATCATCTTACAGA | 798 |
| Dd | | 1714990 | AAAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT | 17149322 |
| OY | | 799 | CACATATATGCAACCGACCAATATATGACGAAGTATATCTGACCAACATACATGTATTTAT | 858 |
| Dd | | 1714931 | CAAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT | 17148727 |
| OY | | 859 | CTTAACATATCTTCTACACACCAATTGCGACCATATATGACACAAATATACACACAA- ---- | 912 |
| Dd | | 1714871 | ATATATATATATATATATATATATATATATATATATATATATATATATATATATAT | 17148121 |
| OY | | 913 | --- AAGAGATATGCTTAAAAAATATCCCAATTTTATTTTGCTGTAAAGGTATCGAAATTAC | 969 |
| Dd | | 1714811 | TATATATATATATATATATATATATATATATATATATATATATATATATATATATAT | 17147522 |
| OY | | 970 | ATCACTCAAAATTTCCCACTGGAATATGACAGAAATTTTACAAATATATCCAAATTTTTAT | 1029 |
| Dd | | 1714751 | ATATATATATATATATATATATATATATATATATATATATATATATATATATATAT | 17146927 |
| OY | | 1030 | ACTCTGCTGCCAAGTTTGTCTAAAGAAATATGCGAAGCATCTCTCTTGATATATAGTCT | 1089 |
| Dd | | 1714651 | AATATACATATATATATATATATATATATATATATATATATATATATATATATATAT | 17146323 |
| OY | | 1090 | TGCTGCTATATATATATATATATATATATATATATATATATATATATATATATATAT | 1149 |
| Dd | | 1714631 | TATATATATATATATATATATATATATATATATATATATATATATATATATATAT | 17145727 |
| OY | | 1150 | AAAAATTA --- ATATACCATGCAAGATATATATGCTGTACAGATGGAGCTGATTAADAA | 1206 |
| Dd | | 1714571 | AAATATATATATATATATATATATATATATATATATATATATATATATATATATAT | 17145122 |
| OY | | 1267 | CATAATTCGAATTCGATGTGGAATATATCGCTGGAAAAAATTTAGGTAAATATGCTGTGC | 1266 |
| Dd | | 1714511 | AAAAAATATATATATATATATATATATATATATATATATATATATATATATATAT | 17144523 |
| OY | | 1267 | TACTGTTTAACTAAAGCATGAATGAATTAACATATATATATATATATATATATATAT | 1326 |
| Dd | | 1714451 | AAATATATATATATATATATATATATATATATATATATATATATATATATATATAT | 17143929 |
| OY | | 1327 | AATAGATCT -TCATCATTTAATGTGTCATCATGTCTTGAGATCTAATATTAACATETA | 1385 |
| Dd | | 1714391 | AATATATATATATATATATATATATATATATATATATATATATATATATATATAT | 17143322 |
| OY | | 1386 | CAACGACATTCACGCAATATGCTTTGGCATGAAATATGCAAAAGCTGGAATATATA | 1445 |
| Dd | | 1714331 | TAA --- AATATATATATATATATATATATATATATATATATATATATATATATAT | 17142757 |
| OY | | 1446 | ATACATTATATGATCACAACCTTAAGCTTTGTTGAAGAACCAATATCTGTAAACCACTGA | 1505 |

[illegible]


```
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See file wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: Patentin version 3.1
SEQ ID NO 17353
LENGTH: 1029
TYPE: DNA
ORGANISM: Clostridium difficile
US-10-282-122A-17353

Query Match      3.7%; Score 175.2; DB 12; Length 1029;
Best Local Similarity 51.6%; Pred. No. 2.4e-18;
Matches 483; Conservative 0; Mismatches 438; Indels 15; Gaps 3;
```

```

QY 994 AGCAGAAAGAAATTTTACATATATATCAATTTTAACTCTCTGCTCCAGTTTGGCTAA 1053
DB 369 CTGTGAAGAAAGAAATTTCCACAAAATTCATATGTAAATATCTATCGGGCCATCTATGCTGA 428
QY 1054 AGAAATTCAGAACATCTTCTCTGTATAGTACTGTGCTGATGAATATGAACCTTGG 1113
DB 429 AGAAGTGCAGAGGATATTTCTTACTACCGTAGAGTGTCTTAAAGATTAAATAGC 488
QY 1114 TGAATCATGTATGAACCAATATAGTATGTCTTAAATATATATACCATTAAGATAT 1173
DB 489 ACAGATGATCAAGATTTGTTTATGAGCCCTTAACTTAAAGATTATACGAATTCACACAT 548
QY 1174 TATAGGTGTACAGATTGGAGCTGCATTTAAAGAACATATTTGCAATTGCATGGAATAT 1233
DB 549 AGTTGGAGTGAACCTTGGAGGACATTTAAATATATATAGCATTTGGAGCTGGATATG 608
QY 1234 CGCTGAAAAAATTTAGTATATATATATATATATATATATATATATATATATATAT 1293
DB 609 TGACGATTTAGGCTATATATATATATATATATATATATATATATATATATATATAT 668
QY 1294 AATTAAGACATATATATATATATATATATATATATATATATATATATATATATAT 1353
DB 669 AATGAGGAGCTTGGAAATTC-----TATGGAGCTATATATATATATATATATAT 722
QY 1354 ATCATGCTTTGAGATCTATATATATATATATATATATATATATATATATATATAT 1413
DB 723 TTGAGTATTTGAGATTTATATATATATATATATATATATATATATATATATATATAT 782
QY 1414 TGGAATGAATATGAAAGGTGAAATATATATATATATATATATATATATATATATATAT 1473
DB 783 TGTATATCTATATGTAAGAGATGAGTTAGAGATCTTTAAAGAGTAAGATGCT 842
QY 1474 TGTGAAGAACAGTATCTGTAAACACTGATATCTATAGCAAAAAAATTAATGTAG 1533
DB 843 TGTAAAGGATATAGCTATCTAGTAAAGATGATGACATGATGTTGCTGAAAAATTAATAG 902
QY 1534 ACTACCATTTGATATCTATATATATATATATATATATATATATATATATATATAT 1569
DB 903 TATGCTATATATATATATATATATATATATATATATATATATATATATATATAT 938

RESULT 10
US-10-282-122A-35500
Sequence 35500. Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
```


PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining prior application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 35500
LENGTH: 996
TYPE: DNA
ORGANISM: Staphylococcus haemolyticus
US-10-282-122A-35500

Query Match 3.6%; Score 170.6; DB 12; Length 996;
Best Local Similarity 49.6%; Pred. No. 1.3e-17;
Matches 468; Conservative 0; Mismatches 469; Indels 6; Gaps 1;

635 AAAAATCTATTTAGTGGAGAGATCTTTGGCAGAGAAATGCAATGACCTGTCAGCA 634
7 AAAATCTATTTAGTGGAGAGATCTTTGGCAGAGAAATGCAATGACCTGTCAGCA 66
695 CATGATATGATGATTAATGATGAGAGATCTTTGGCAGAGAAATGCAATGACCT 754
67 AATGACATGATGATTTAATGATGAGAGAAATGCAATGATGAGAAATGAT 126
755 TACGAAAAATTTAAATATTTTACCCATATCATCTACAGCAACATATATGCAAC 814
127 CACCATCAAAATTAACGTTATTTAAAGATGCTTTAGATGATTAATTAAGCACT 186
815 AGCAATATAGAGAGATTTATCTGACACAAATACATGATTTATCTTAATCTCTACA 874
187 AGTATATTAAGAGAGAGAGAGATTTTATGATATATTTAATGATGATGATGAT 246
875 CACCAATAGAGAGATTTATCTGACACAAATACATGATTTATCTTAATCTCTACA 934
247 GCAATGAGAGAGATTTATCTGACACAAATACATGATTTATCTTAATCTCTACA 306
935 CCAATATTAATTTAGTGGAGAGATCTTTGGCAGAGAAATGCAATGACCTGTCAGCA 994
307 CATGATGAGAGATTTATGATGAGAGATCTTTGGCAGAGAAATGCAATGACCTGTCAGCA 366
995 GCAAGAGAGATTTATGATGAGAGATCTTTGGCAGAGAAATGCAATGACCTGTCAGCA 1054
367 TCAATTTAGAGAGATTTATGATGAGAGATCTTTGGCAGAGAAATGCAATGACCTGTCAGCA 426
1055 GAAATTTAGAGAGATTTATGATGAGAGATCTTTGGCAGAGAAATGCAATGACCTGTCAGCA 1114
427 GATGATGAGAGATTTATGATGAGAGATCTTTGGCAGAGAAATGCAATGACCTGTCAGCA 486
1115 GATGATGAGAGATTTATGATGAGAGATCTTTGGCAGAGAAATGCAATGACCTGTCAGCA 1174
487 AATGATGAGAGATTTATGATGAGAGATCTTTGGCAGAGAAATGCAATGACCTGTCAGCA 546
1175 AATGATGAGAGATTTATGATGAGAGATCTTTGGCAGAGAAATGCAATGACCTGTCAGCA 1234
547 GTCGATGAGAGATTTATGATGAGAGATCTTTGGCAGAGAAATGCAATGACCTGTCAGCA 606
1235 GTCGATGAGAGATTTATGATGAGAGATCTTTGGCAGAGAAATGCAATGACCTGTCAGCA 1294
607 GATGATGAGAGATTTATGATGAGAGATCTTTGGCAGAGAAATGCAATGACCTGTCAGCA 666
1295 AATGATGAGAGATTTATGATGAGAGATCTTTGGCAGAGAAATGCAATGACCTGTCAGCA 1354
667 AATGATGAGAGATTTATGATGAGAGATCTTTGGCAGAGAAATGCAATGACCTGTCAGCA 720
1355 TCAATTTAGAGAGATTTATGATGAGAGATCTTTGGCAGAGAAATGCAATGACCTGTCAGCA 1414

DB 721 GATGATGAGAGATTTATGATGAGAGATCTTTGGCAGAGAAATGCAATGACCTGTCAGCA 780
QY 1415 GATGATGAGAGATTTATGATGAGAGATCTTTGGCAGAGAAATGCAATGACCTGTCAGCA 1474
DB 781 GATGATGAGAGATTTATGATGAGAGATCTTTGGCAGAGAAATGCAATGACCTGTCAGCA 840
QY 1475 GATGATGAGAGATTTATGATGAGAGATCTTTGGCAGAGAAATGCAATGACCTGTCAGCA 1534
DB 841 GATGATGAGAGATTTATGATGAGAGATCTTTGGCAGAGAAATGCAATGACCTGTCAGCA 900
QY 1535 GATGATGAGAGATTTATGATGAGAGATCTTTGGCAGAGAAATGCAATGACCTGTCAGCA 1594
DB 901 GATGATGAGAGATTTATGATGAGAGATCTTTGGCAGAGAAATGCAATGACCTGTCAGCA 964

RESULT 11

US-09-815-242-8974
Sequence 8974, Application US/09815242

Patent No. US2002061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Kari L.

APPLICANT: Zyskind, Judith W.

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John D.

APPLICANT: Carr, Grant U.

APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in

FILE REFERENCE: ELITRA 011A

CURRENT APPLICATION NUMBER: US/09/815,242

PRIOR FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO: 8974

LENGTH: 999

TYPE: DNA

ORGANISM: Staphylococcus aureus

FEATURE:

NAME/KEY: CDS

LOCATION: (1)...(999)

US-09-815-242-8974

Query Match 3.5%; Score 168.6; DB 9; Length 999;
Best Local Similarity 51.4%; Pred. No. 2.6e-17;
Matches 505; Conservative 0; Mismatches 454; Indels 24; Gaps 4;

635 AAAATGATTTTATGATGAGAGATCTTTGGCAGAGAAATGCAATGACCTGTCAGCA 694
7 AAAATGATTTTATGATGAGAGATCTTTGGCAGAGAAATGCAATGACCTGTCAGCA 66
695 CATGATATGATGATTAATGATGAGAGATCTTTGGCAGAGAAATGCAATGACCTGTCAGCA 754
67 AATGACATGATGATTTAATGATGAGAGATCTTTGGCAGAGAAATGCAATGATGAGAAATGAT 126
755 TACGAAAAATTTAAATATTTTACCCATATCATCTACAGCAACATATATGCAAC 814
127 TGTATATCAAAATTAAGATTTAAATATGAGAGATCTTTGGCAGAGAAATGCAATGACCTGTCAGCA 186

QY 815 AGCAATATAGCAAGATATATATGCAACATACATGATATATCTTAATCTTCA 874
 DB 187 TCAGATATACCAAGACATTCATTTGAGATATTTACTTAATG---CTTTACCTTCT 243
 QY 875 CACACATTTAGCAACATATGTAACAATATACACACAAACAGCATATGTGTAATAATCT 934
 DB 244 AAGAGCATCCGAAAGTTCCTTCATAATATATGTAAGC-----TACCTCTAAAAAG 297
 QY 935 CCAATATATTTGTAAGGTAATCGAAATTAATCATCTCAATTTCCAGTAATA 994
 DB 298 ACTTTATACATGCTGTAAGGTATGAAATGGAGCCTTTAAACGTGTCAAAATG 357
 QY 995 GCAGAAATATTT-----ACAAATTAATCCAAATTTTAACTCTGTGTCAAGT 1045
 DB 358 ATTGAAGATTTCTATTTCACTTAATATATATGAGGATTTGGCGTGTCTGAGGGCAAGT 417
 QY 1046 TTTCCTAAAGAAATGCAAGATCTCTCTGTAGATATGATCTGTGGATATAA 1105
 DB 418 CATGCGAAGAAATTTAGTCAAGCACTACCTGCTGCTTCACTCAAAAGATATA 477
 QY 1106 GAACTTGTAATCATTTGATAGAAACAAATAGTAATGATTTTAAAAATATATACAT 1165
 DB 478 AGGTATAGTAATTAACGCAAGATTTATTTATGAAATGATTTTTCGTGTGACAGAAAT 537
 QY 1166 CAAATATATATAGGTATACAGATGAGCTGATTAAGAAATATATGCAATTTGCAATGAT 1225
 DB 538 GATGATCTTATGTTGTGTAAGTGTGATGATGAAATATATATGCAAGTACAGT 597
 QY 1226 GGAATATGCTGCAAAAAATTTAGGTAATATGCTGTCTGATCTTAATCTAAAGGC 1285
 DB 598 GGTATGTAAGTGTGATGCTGATACGCTGATATGCAAAAGCTGATATATGCTGTGCG 657
 QY 1286 ATCAATGAAATTTAAACATATATATAGCAAAAAATCATTCATATGATCTTATCATATA 1345
 DB 658 TTAGCGGAATTAAGTGA-----TTAGGTGAAAGTATAGTGCCTGATCTATGCAATTT 711
 QY 1346 ATGTGTCATCATGCTTGTGAGATCTAATATTAACATGATACAGCAACATTCACGCAAT 1405
 DB 712 CTAGTTTATAGTGTATCGGTGATCTTAATGCTTATCTGCACTCAACATTTCTGGAAT 771
 QY 1406 ATGCTTTTGGACTAGAAATAGGAAAGGTAGAAATATTAATACCTTAATATGATCAAC 1465
 DB 772 TTCACTTATGAAATTAATTAAGTGTGACAAAGTGAATCAATGACAAAGCATTTATCTAAATG 831
 QY 1466 CTAAAGCTTGTGAAAGCAACAGTACTGTAAACCACTGATCATTTAGCAAAAAAACTT 1525
 DB 832 AATATGTTTGTAGAGTATTTATACACTTAATCACTTTATCATTAGCTTAAGAAAA 891
 QY 1526 AATGTAGAACTACCAATTTGCAATTTATTAATTTATATGAGAAATATATCACTA 1585
 DB 892 AATGTAGAAATGCAATTTCAATGATATATAGATATTTTGAATAATATCTCAGTA 951
 QY 1586 GATTAAGCCATATCAAAACATATT 1608
 DB 952 AAAGAAATGCTTAAGATTTAAT 974

APPLICANT: Forsyth, R.
 APPLICANT: Xu, H.
 TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 FILE REFERENCE: ELITRA 034
 CURRENT APPLICATION NUMBER: US/10/282,122A
 CURRENT FILING DATE: 2003-02-20
 PRIOR APPLICATION NUMBER: 60/191,078
 PRIOR FILING DATE: 2000-03-21
 PRIOR APPLICATION NUMBER: 60/206,848
 PRIOR FILING DATE: 2000-05-23
 PRIOR APPLICATION NUMBER: 60/207,727
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: 60/230,335
 PRIOR FILING DATE: 2000-09-06
 PRIOR APPLICATION NUMBER: 60/230,347
 PRIOR FILING DATE: 2000-09-09
 PRIOR APPLICATION NUMBER: 60/242,578
 PRIOR FILING DATE: 2000-10-23
 PRIOR APPLICATION NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/257,931
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 60/267,636
 PRIOR FILING DATE: 2001-02-09
 PRIOR APPLICATION NUMBER: 60/269,308
 PRIOR FILING DATE: 2001-02-16
 Remaining Prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 78614
 SOFTWARE: Patent version 3.1
 SEQ ID NO 15395
 LENGTH: 999
 TYPE: DNA
 ORGANISM: Clostridium acetobutylicum
 US-10-282-122A-15395

Query Match 3.5%; Score 168.2; DB 12; Length 999;
 Best Local Similarity 51.0%; Pred. No. 3.1e-17;
 Matches 482; Conservative 0; Mismatches 448; Indels 15; Gaps 3;

QY 648 TAGGTGAGATCATTTGGCAGACAGATAGCAATTCATGCTGACAGCATGTATATCAG 707
 DB 20 TAGGAGAGAGAACTTTGGTACAGCATTCGATATATGCTGTAAAGGATCAATAG 79
 QY 708 TTAATTTAGGAGCGTATCATGAATATATATATATATATATATATATATATATAT 767
 DB 80 TTGTGATTTGGATAGAAATTAAGAAATCTTGAAGATTAATATATATATATATATAT 139
 QY 768 TAAATATTTACCAATATCATCTACAGACAAATATATGCAACAGCATATATAGAG 827
 DB 140 CAAATATCTTCAACAAATATATATATATATATATATATATATATATATATATATAT 199
 QY 828 AAGTATATCTGCAACAAATATATATATATATATATATATATATATATATATATAT 887
 DB 200 AGG---CAGCTAAGAAAGCAAAATATATATATATATATATATATATATATATAT 256
 QY 888 CCAATGTACAAATATCAACAAAGCAATATATATATATATATATATATATATATATAT 947
 DB 257 AAGTATCAGAAATATATATATATATATATATATATATATATATATATATATATAT 310
 QY 948 GTAATAGGTATCGAAATATATATATATATATATATATATATATATATATATATAT 1007
 DB 311 TTGCAAAAGCATAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 370
 QY 1008 TACAAATATATCAATTTTATATATATATATATATATATATATATATATATATATAT 1067
 DB 371 TTTATATATATCTGTTGTATATATATATATATATATATATATATATATATATATAT 430
 QY 1068 ATCTTCTGTATATATATATATATATATATATATATATATATATATATATATATAT 1127
 DB 431 ATATTCACATACAGTATGTTATATATATATATATATATATATATATATATATATAT 490
 QY 1128 AAACATATATATATATATATATATATATATATATATATATATATATATATATATAT 1187

Db 491 ATGCTTTTATGACTACAGCTTTAGGCTGTACACAAATAGTATTTAGAGCGTTGAAA 550
Qy 1188 TTGGAGCTGCATTTAAAGAAATTTGCAATTCATGGAATTAATCGTGGAAAAATTT 1247
Db 551 TAGGAGGAGCAGTAAATATATATGCTCTGATGCAATGATGATGGAATAGGTT 610
Qy 1248 TAGGTATATGCTGTGTTGCTACTGTTATATACCTAAAGGATGATGAATTAACACATAT 1307
Db 611 ATGCTGACACACAAAGAGCGCTTATGACAAAGGATGATGAATTAATGAGA---- 666
Qy 1308 ATATAGCAAAAAATCATTCATTAATGATCTTCATATATATGCTCATCATGCTTGGAG 1367
Db 667 --ATAGCGCTTAAAGCTTGGAGGTAAGCCGAAACTTTTGGACCTTACAGGCAATGGTG 724
Qy 1368 ATCTATATTAATCATGTACACAGAACTTCACGAAATATGCTTTGGAATGAAATAG 1427
Db 725 ACCCTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 784
Qy 1428 GAAAGGTGAAATATTAATATATATATATATATATATATATATATATATATATATAT 1487
Db 785 GTAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 844
Qy 1488 GTAGCTTAATACCTGATATATATATATATATATATATATATATATATATATATAT 1547
Db 845 AAGCGTGTACACCTTTATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 904
Qy 1548 TATCTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1592
Db 905 CTTCACCTTATTAAGGTTCTATTTGGAATGAGATCTTAATAAAG 949

RESULT 13
US-10-282-122A-34158
Sequence 34158, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:

APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 34158
LENGTH: 999
TYPE: DNA
ORGANISM: Staphylococcus aureus
US-10-282-122A-34158

Query Match 3.5%; Score 165.4; DB 12; Length 999;
Best Local Similarity 51.2%; Pred. No. 8.5e-17;
Matches 503; Conservative 0; Mismatches 456; Indels 24; Gaps 4;

Qy 635 AAATACGATATTTAGGTCAGATCATTTGGACAGCAATATGCAATTCAGTCATGACGA 694
Db 7 AAATATACGATTTTGGTATGGAATTTTGGACAGCCCTTCCCAATTTCTTGACAGA 66
Qy 695 CATGATATACGATTAATCTTAATGGGAGCTGATCATGAAATTTACATATTAACACT 754
Db 67 AATGACATGATGTTTGGATGAGGGTAAATAATCAAGATGCTGTTGATGATTAATTA 126
Qy 755 TACGAAAAATTTAAATATTTATACCATATATATATATATATATATATATATATATAT 814
Db 127 TGTATACAAATTAATTAATTTAAATACGAAATTTAGATTTACATCATCGTACT 186
Qy 815 AGCAATATAGACGAGTATTTCTGACAAATATATATATATATATATATATATATATAT 874
Db 187 TCAGATATAGACCAAAATTCATATTTGCAATATTTACTTATATG--CTTACCTACT 243
Qy 875 CACCAATTTGCGCAGCATATGTCACAAATATCAACAAACAGATATGTTAAATACT 934
Db 244 AATGCAATGCGAAGATTTGCTACTCAATTAATGATAC--TGACCTTAATAAG 297
Qy 935 CCAATATATTTGATTAAGATATGAAATTAATATATATATATATATATATATATATAT 994
Db 298 ACTTTATATATATGCTTAAGATATGAAATGAGGAGCTTAAACGTGTGACAGAAATG 357
Qy 995 GCAGAAATTTT-----ACATATATCCATTTTATATCTCTGTGCTCAAT 1045
Db 358 ATGGAATTTCTATTTACCCGAAATTAATGAGATTTGGCGTGTGACAGGCTCAAT 417
Qy 1046 TTGCTAAAGAAATTCAGAACATCTTCTGTATATATATATATATATATATATATAT 1105
Db 418 CATGCGAAGATTTGATGATCAAGCAACCAATGATGCTGCAATCATCAAAAGTAA 477
Qy 1106 GAACTGTGATCATTTGATGAAACATTAATGATGATTTCTTAATAATATATATAT 1165
Db 478 AGTGAATTAATTAATGCAAGATTTATTAATGATGATTTGGTGTGACAGAAAT 537
Qy 1166 CAGATATTAATAGGTATCAGATTTGAGAGCTGATTAAGACATTAATTCATATGATGT 1225
Db 538 GATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 597
Qy 1226 GGAATATATGCTGGAATAATTTAGTAAATATGCTGTGCTACTGTTTAACTAAAGGC 1285
Db 598 GGAATGATGCTGGAATTTGCTACGCTGATTAATGCAAAAGCTGCTTAATGATGCTGTGT 657
Qy 1286 ATGAATGAATTAATTAATCAATATATATATATATATATATATATATATATATATAT 1345
Db 658 TTAGCGGAATTTGATGAA-----TTAGTGAATTTAGTGGCGGATCTATGACATTT 711
Qy 1346 ATTTGTCATCATGCTTTGGAATCTTAATTAATTAATTAATTAATTAATTAATTAAT 1405
Db 712 CTAGGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 771
Qy 1406 ATGCTTTTGAATGAATTAAGAAAGGTAGAAATATTAATTAATTAATTAATTAATTAAT 1465
Db 772 TTCAATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 831
Qy 1466 CTAAAGCTTTGTAAGAAACGATGATTAATTAATTAATTAATTAATTAATTAATTAAT 1525
Db 832 AATATGTTGTTGAAGATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 891
Qy 1526 AATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1585

QY 1302 CACTATATATAGCAAAAAATCTATCATATGATCTTCATACATTAATGGTCCATCATGTC 1361
DB 725 GTCTGGGTGTCGATGGGT-----GCCAATCCCTGACATTTTCGTTGAGTGCTG 778
QY 1362 TTGGAGATCTAATATTAACATGATACACAGAACATTCACGCAATATGGCTTTGGACTAG 1421
DB 779 TGGGGGACTGATCGTACTTGTACAGAGTTCATTCGCAATGGGGTCCGCTACT 838
QY 1422 AATAGGAAAAGGTAGAAATTAATTAATCTTATATGATCAACCTTAAGCTTTGAG 1481
DB 839 TGTAGGACAGGACATATGCTGAGAGAGTTCTGAGAAAACATGGGAATGGCTTGAGG 898
QY 1482 GAACAGTACTGTAAACCACTGATATCTATAGCAAAAAAATTAATAGTAGACACCA 1541
DB 899 GTGTGCAACGACAAAGACAGACGATGAGCTAGCTCAGCTAGGTGTAGATGCCA 958
QY 1542 TTTCATATCTATTTACAAATTTATACATGAGATATATCACTAGATTAAGCCATATCA 1601
DB 959 TTAATCAAACTATATATGTTTGTATATGAGAGATATCAAAAAAGCAGCAAAAG 1018
QY 1602 ACATATAT 1610
DB 1019 AATCATGT 1027

RESULT 16
US-10-311-455-1369/c
; Sequence 1369, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1369
; LENGTH: 5641
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1369

Query Match 3.3%; Score 155; DB 14; Length 5641;
Best Local Similarity 43.9%; Pred. No. 7,4e-15;
Matches 1027; Conservative 0; Mismatches 1280; Indels 34; Gaps 7;

QY 176 ATTTTCTCAATCTCCAAATTTTAACATGACAGATTTTAATGTAACAGACACCTG 235
DB 4946 ATTTAATTAATAACATTAATAATAACAAATCAACATTTTCTTAATAATTAATATCTTA 4887
QY 236 TACTCTTATGACAGACCTCTATAAATTATCAAGGCTCTTATACATAGAAAACACTGTAT 295
DB 4886 CACCCCAATTTTATATTTTAAATAATAATAATAATAATAATAATAATAATAATAATA 4830
QY 296 TAAAGAAATACCTCTCTCTGTCATCCATCCATATTTTCCATTTATTTTCAATATCATGA 355
DB 4829 AGGATATATTTTCACTACTTCAAAACTTAAATACTCAATATTTAAATAATAATAATA 4770
QY 356 GTTCTTATGACATACATATTAATAATCAATCAAAACAAATAATATAGTACTATCTAAT 415
DB 4769 TAAATCTATTTTCAATCTATCACTATCACTATTTCTTAATAATAATAATAATAATA 4710

QY 416 AATTAATAATCTCTCAATAGT---CTTAGCAGAGACTACAAATTAATGAGACCTTA 471
DB 4709 AATTAATATATATATATATTAATAATAATAATAATAATAATAATAATAATAATAATA 4650
QY 472 ACTTATTAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 531
DB 4649 AAAAAAAAAAAAAAAAAATTAATTAATAATAATAATAATAATAATAATAATAATAATA 4590
QY 532 TGACATGATTAACATATGATACATCTTTTAACACTAGATATATATATATATAT 591
DB 4589 ATATATCTATCTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 4530
QY 592 ATATATTTGACCTGCTTCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 651
DB 4529 ATATCTTAATAATCAATCTTCCCTACTAATTAATAATAATAATAATAATAATAATAATA 4470
QY 652 TGCAGATCAATTTGGACAGCAATAGCAATGCACTGTCAGACATGATATCACTTAA 711
DB 4469 TAAACCTCAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 4410
QY 712 CTATAGGAGACGATCATAGAAATTTACATATTAACATTAACCTTACCGAAAAATTTAA 771
DB 4409 ATATAT-----AAATTAATAATAATAATAATAATAATAATAATAATAATAATAATA 4357
QY 772 ATATTTACCCATATATCATCTACAGACAAATATATGACACGATATAGACGAGT 831
DB 4356 AATCAATTAACCTCAAAATATATACCTATATATATATATATTAACGAAAAATCAAAAAA 4297
QY 832 ATATATGACAAATATACATGATATATCTTAATCTATCTTACATTAATTAATTAATTAAT 891
DB 4296 ATATTAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 4237
QY 892 ATGTACAAATTAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 951
DB 4236 ATTAACAAATTAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 4177
QY 952 TAAAGATTCGAAATTAATACATCACTCAATTTCCAGTGAATATGACGAAATTTTAA 1011
DB 4176 TAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 4117
QY 1012 ATATATCAATTTTATATCTCTCTGTCACAGTTTGTGTAAGAAATGACGAAATCT 1071
DB 4116 CTTTAATTAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 4057
QY 1072 TCCTGTATATATGATCTGCTGCTGATATTAAGAACTTGGAATCTTGTATGAAC 1131
DB 4056 AATCTATCTAATTTTCACTCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3997
QY 1132 AATAGTAATGATGTTCTAATAATAATAATAATAATAATAATAATAATAATAATAATA 1191
DB 3996 AAAAAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 3937
QY 1192 AGCTGATTAAGACATTAATGCAATTCGATGGAATTAATCGCTGGAATAATTTAG 1251
DB 3936 TAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 3877
QY 1252 TAAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1311
DB 3876 AAAAAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 3817
QY 1312 AGCAAAAAATCAATTAATGATCTTCAATTAATTAATTAATTAATTAATTAATTAAT 1371
DB 3816 AATTAACAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 3767
QY 1372 AATATTAACATGACAGACAGACATTCAGCAATATGCTTTGACATGAATATGAAA 1431
DB 3766 ATATATTAATTTGAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3707
QY 1432 AGGTAGAAATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1491
DB 3706 AATCATATCAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 3651

QY 1442 ATAAATACATTAATGATGACAACTAAAGCTTTGTAAGGACGATCTGTAACCA 1501
DB 318 ATGGATCAAGCATTTATCTGAAATGATATGATGTTGTAAGGATTTATTAACAATAACA 259
QY 1502 CTGATATCATTTGCAAAAAAACTTAATGAGAACTACCAATTTGATATCTATTACAT 1561
DB 258 GTTATATCATTTAGCTTAAGAAAAAATGTGATATGCCAATTACAAATGATATATAGA 199
QY 1562 TTATTCATGAGAAATATATCTAGATTAAGGCAATATCAACATTTATCTAGCTATC 1621
DB 198 GATATTTTGAATAATCTCAGTTAAAGATGCGTAAAGATTTATGAGCGCATAAA 139
QY 1622 ATCTTTGATTAATTTCAACAAATGATTAATAAGGTCGCAAAAT 1666
DB 138 AAATCTGAATTAATGTGTATTTATTTACATATATGCGCTAATAT 94

RESULT 18
US-10-329-624-513/c
; Sequence 513, Application US/10329624
; Publication No. US20040063037A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; GIL H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Rannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/329,624
; FILING DATE: 27-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/956,171
; FILING DATE: October 20, 1997
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 513:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 918 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 513:
US-10-329-624-513

Query Match 3.2%; Score 154.2; DB 12; Length 918;
Best Local Similarity 53.6%; Pred. No. 4.9e-15;
Matches 346; Conservative 0; Mismatches 293; Indels 6; Gaps 1;

QY 1022 ATTTTATACCTCTGCTGCAAGTTTGGCTAAAGAAATGCAACATCTCTCTAGT 1081
DB 732 ATGGGCGTGTTCAGGGCCAGTCTATGCGAAGAGTTGTGTCAAGAACCACTACA 673
QY 1082 ATAGTACTGCTGCTGATATAAAGAACTGGGATTCATTTGATAGAAACAAATAGTAAT 1141
DB 672 GTTGTGCTTCATCAAAAAGATTAAGTAAAGTAAGTAATTAACGAAATTTATTTAGAT 613
QY 1142 GATGTTCTAAAAATTAATATATCAATCAAGATTTATATGTTGATACAGATTTGAGCTGATTA 1201
DB 612 GATTTATTTGGTGTGACACGAATGATGATCTGATTTGGTGTGAACCTTGCTGTCATTG 553
QY 1202 AAGACATTAATTCGAATTTGCAATGATGAAATATGCTGCAAAAAATTTAGTAAATATGCT 1261
DB 552 AAAAAATTCATTCGCAATGACAGATGATGATGATGATGATGATGATGATGATGATGATGAT 493
QY 1262 GTTGCTACTGTTTAACTAAAGGATGAATGAATTTAAACACTATATATAGCAAAAAT 1321
DB 492 AAAGCTGATTAATGATCTGCTGCTTACGGAATTTAGTGA-----TTAGTGAAAAAG 439
QY 1322 CATTAATGATGATCTGATCATTAATTTGGTCCATCATGCTTGGAGATCTAATATTAACA 1381
DB 438 TTAGTGCGGATCTCTATGACATTTCTAGGTTTATGATGATGATGATGATGATGATGATGAT 379
QY 1382 TGTAACAAGACATTTACGCAATATGCGTTTGAATGAAATAGCAAAAGGTAGAAAT 1441
DB 378 TGCAATCAACACATTTCTGGAATTTACATATGATATATACTTGGAACAGGTGATCA 319
QY 1442 ATAAATATTAATTAATATATCAACAAAGCTTGTGAAGGAACGATCTGTAACCA 1501
DB 318 ATGATCAACATTTATCTGAAAGAAATGATGTTGTTGAAAGATTTATATCAACTAAATCA 259
QY 1502 CTGATATCATTTGCAAAAAAACTTAATGAGAACTAACCAATTTGATATCTATTACAT 1561
DB 258 GTTATCATTTAGCTTAAGAAAAAATGTGATATGCCAATTAACAATCATTAATATAGA 199
QY 1562 TTATTCATGAGAAATATATCTAGATTAAGGCAATATCAACATTTATCTAGCTATC 1621
DB 198 GATATTTTGAATAATCTCAGTTAAAGATGCGTAAAGATTTATGAGCGCATAAA 139
QY 1622 ATCTTTGATTAATTTCAACAAATGATTAATAAGGTCGCAAAAT 1666
DB 138 AAATCTGAATTAATGTGTATTTATTTACATATATGCGCTAATAT 94

RESULT 19
US-10-311-455-2128/c
; Sequence 2128, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: Olek, Alexander
; APPLICANT: PREPNBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Deter
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 2128
; LENGTH: 15548
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-2128

Query Match 3.2%; Score 153; DB 14; Length 15548;
Best Local Similarity 44.7%; Pred. No. 2.3e-14;
Matches 981; Conservative 0; Mismatches 1185; Indels 27; Gaps 9;

QY 237 ACTCTTACGACAGACCTTAACTATACAGAGCTCTTACATCAAAACAACTGAT 296
DB 13534 ACGCATACGATTAATAATATATATTAACACATACGATTAATAATATATATAT 13475
QY 237 CAAGAATACCTCTCTCTGTCATCATATATTTCCAAATTATTTCAATATCATAG 356
DB 13474 CATATATAACATACGATTAATAATATATATATTAATAATATATATATTA 13416
QY 357 TTCTTATTCGACATACATATTAATATCAACAAATAATAGTACTATCTATTA 416
DB 13415 ATATATAATATATATATATATATATATATATATATATATATATATAT 13356
QY 417 ATAAAAAATCCCTAATAGCTTTGACAGAGAACTACAAATATATATAGACATTA 476
DB 13355 ATATATAATATACGCTAATAATATATATATATATATATATATATATATAT 13296
QY 477 TTATATATTTTCAATTTATATATATCTCTCATATTTTGCATATATATCAT 536
DB 13295 CTAAATATATATATATATATATATATATATATATATATATATATATAT 13236
QY 537 TGATTAATATATATATATATATATATATATATATATATATATATATATAT 596
DB 13235 TATATATATATATATATATATATATATATATATATATATATATATATAT 13177
QY 597 TTTGACCTGTTTCTGATATATATATATATATATATATATATATATATAT 656
DB 13176 ATATATATAATATATATATATATATATATATATATATATATATATAT 13117
QY 657 GATCTTTTGGACAGCAATAGCAATGCTGACAGCAATGATATATATATATAT 716
DB 13116 ATAAAAATATATATATATATATATATATATATATATATATATATAT 13057
QY 717 GGGGACGTATATATATATATATATATATATATATATATATATATATAT 776
DB 13056 TATATATATATATATATATATATATATATATATATATATATATATAT 12997
QY 777 TACCCATATATATATATATATATATATATATATATATATATATATATAT 836
DB 12996 TATATATATATATATATATATATATATATATATATATATATATATAT 12937
QY 837 CTGACATATATATATATATATATATATATATATATATATATATATATAT 896
DB 12936 ATATATATATATATATATATATATATATATATATATATATATATATAT 12878
QY 897 CACAAATATATATATATATATATATATATATATATATATATATATATAT 956
DB 12877 TACGAAATATATATATATATATATATATATATATATATATATATAT 12818
QY 957 GTATCGAATATATATATATATATATATATATATATATATATATATATAT 1016
DB 12817 ATATATATATATATATATATATATATATATATATATATATATATAT 12758
QY 1017 ATCCATTTTATATCTCTGCTGCAATTTTGTCTTAAAGAAATGCGAATCT 1076
DB 12757 TATATATATATATATATATATATATATATATATATATATATATATAT 12698
QY 1077 GTAGATATATATCTCTGCTGATATATATATATATATATATATATATAT 1136
DB 12697 TATATATATATATATATATATATATATATATATATATATATATATAT 12638
QY 1137 GTATGATGTTCTTAAATA-----ATATACATCAAGATTTTATAGTGAC 1188
DB 12637 ATATATATATATATATATATATATATATATATATATATATATATAT 12578
QY 1189 TGAAGCTCATTAAGAACTATATATATATATATATATATATATATATATAT 1248
DB 12577 ATAAAAATATATATATATATATATATATATATATATATATATATAT 12516
QY 1249 AGGTATATATCTGTTGCTACTGTTATATATATATATATATATATATAT 1308

DB 12517 AATATATATATATATATATATATATATATATATATATATATATATAT 12458
QY 1309 TATGCAAAAAATCATTCATATATATATATATATATATATATATATATAT 1368
DB 12457 TATGCAAAAAATCATTCATATATATATATATATATATATATATATAT 12398
QY 1369 TCTAATATATATATATATATATATATATATATATATATATATATAT 1428
DB 12397 CATATATATATATATATATATATATATATATATATATATATATAT 12338
QY 1429 AAAAGGTGAATATATATATATATATATATATATATATATATATATATAT 1487
DB 12337 TAAATATATATATATATATATATATATATATATATATATATATAT 12278
QY 1488 GTACTGTAAACCACTGATATATATATATATATATATATATATATAT 1546
DB 12277 ATATATATATATATATATATATATATATATATATATATATATAT 12218
QY 1547 ATATGATTTTACATTTATATATATATATATATATATATATATATATAT 1606
DB 12217 ATATATATATATATATATATATATATATATATATATATATATAT 12158
QY 1607 TTATCTTATGCTATATATATATATATATATATATATATATATATAT 1666
DB 12157 CTATATAAATATATATATATATATATATATATATATATATATATAT 12098
QY 1667 ATGAAATATATATATATATATATATATATATATATATATATATATAT 1726
DB 12097 ATATATATATATATATATATATATATATATATATATATATATATAT 12042
QY 1727 ACTAAGTATATATATATATATATATATATATATATATATATATATAT 1786
DB 12041 CATATATAAATATATATATATATATATATATATATATATATATATAT 11982
QY 1787 TATATATAAATATATATATATATATATATATATATATATATATATAT 1846
DB 11981 ACATATATAAATATATATATATATATATATATATATATATATATAT 11922
QY 1847 GTCAATATATATATATATATATATATATATATATATATATATATAT 1906
DB 11921 ATAAAAATATATATATATATATATATATATATATATATATATATAT 11862
QY 1907 ATGATATATATATATATATATATATATATATATATATATATATATAT 1966
DB 11861 AATATAAATATATATATATATATATATATATATATATATATATAT 11802
QY 1967 TAAATATATATATATATATATATATATATATATATATATATATATAT 2026
DB 11801 AAAATATATATATATATATATATATATATATATATATATATATATAT 11742
QY 2027 TCTATATATATATATATATATATATATATATATATATATATATATAT 2086
DB 11741 AATATATAAATATATATATATATATATATATATATATATATATATAT 11682
QY 2087 AGTCTTATATATATATATATATATATATATATATATATATATATATAT 2146
DB 11681 AATATATAAATATATATATATATATATATATATATATATATATATAT 11622
QY 2147 TGAAGATATATATATATATATATATATATATATATATATATATATAT 2206
DB 11621 ATAAAAATATATATATATATATATATATATATATATATATATATAT 11562
QY 2207 TAAACATATATATATATATATATATATATATATATATATATATATAT 2266
DB 11561 TAAAAATATATATATATATATATATATATATATATATATATATATAT 11506
QY 2267 ATATATAAAGAAATCGTATATATATATATATATATATATATATATAT 2326
DB 11505 AATATAAATATATATATATATATATATATATATATATATATATATAT 11452
QY 2327 AAAATCAAGTCTATATATATATATATATATATATATATATATATATAT 2386

QY 1167 AAGATATTATAGGTGACAGATTGGAGCTGATTAAGACATTAATTGCAATGTCATGTG 1226
DB 542 ATGATGTGATTTGGTGTGAAACAGGGCGGCTTAATAAATATTATTCCTTGGCGCGCG 601
QY 1227 GAATTAATCGCTGGAAAAAATTAGTATAATATCTGCTGCTAGCTTAATACTAAAGCA 1286
DB 602 GAGCGATTTCATGGCTTGGCTTGGTGTATATGCAAGCGCAATTATATGACTCGGTGTT 661
QY 1287 TGAATGAATTAACACTATATATAGCAAAAAATCTTCAATAGATCTTCAATCTTAA 1346
DB 662 TAGCAAAATTTAGTCGTTTGGGCTTGCATGGG-----AGCAAACTTAACTTAA 715
QY 1347 TTGGTTCATCATGCTTGGAGATCTAATTAATTAACATGACAGAACTTACAGCAATA 1406
DB 716 TTGCTTTAAGTGTGTTGGGGAATTAATGTTTACGTTGACAGATGTCATTCGAAAT 775
QY 1407 TTGCTTTGACACTGAATATGAAAAAGTAAATATTAATATACATTAATATGATCAAC 1466
DB 776 GGGCTGGCGGTAATTTATTTAGAAAAAGCCATTAAGTAAAGTCTTGAAGAAACATGG 835
QY 1467 TAAAGCTTGTGAAGAACAGTACTGTAAACCACTGATATCATTAGCAAAAAAATTA 1526
DB 836 GCATGATTTGGAGAGTGTTCACAACTAAAGAGCTATGAATTTGGCAACAACTAG 895
QY 1527 ATGTAGAACTACCAATTTGCAATCTTATTTCAATTTTATTAATGAGATATATCACTAG 1586
DB 896 AAGTGAATGCCATACAGAAACAATCTAACATGCTTGTACAAATGAGAAATGTAC 955
QY 1587 ATAAAGCATATCAACATATTAAT 1610
DB 956 AACAGACAGCAAAAGAAATCATGT 979

RESULT 22

US-09-815-242-6755
Sequence 6755, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Traxick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
PRIORITY FILING DATE: 2001-03-21
PRIORITY FILING DATE: 60/191,078
PRIORITY FILING DATE: 2000-03-21
PRIORITY FILING DATE: 60/206,848
PRIORITY FILING DATE: 2000-05-23
PRIORITY FILING DATE: 60/207,727
PRIORITY FILING DATE: 2000-05-26
PRIORITY FILING DATE: 60/242,578
PRIORITY FILING DATE: 2000-10-23
PRIORITY FILING DATE: 60/253,625
PRIORITY FILING DATE: 2000-11-27
PRIORITY FILING DATE: 60/257,931
PRIORITY FILING DATE: 2000-12-22
PRIORITY FILING DATE: 60/269,308
PRIORITY FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6755
LENGTH: 1023
TYPE: DNA
ORGANISM: Enterococcus faecalis
FEATURE:

NAME/KEY: CDS
LOCATION: (1) ... (1023)
US-09-815-242-6755
Query Match 3.1%; Score 148.4; DB 9; Length 1023;
Best Local Similarity 49.1%; Pred. No. 4,36-14;
Matches 483; Conservative 0; Mismatches 491; Indels 10; Gaps 3;
QY 629 AAAATGAATCAGATTTTGGTGCAGGATCATTTGGACAGCAATAGCAATTCGCTG 688
DB 4 AACCAAAAGTCCGCTTTTGGTCCAGTCTTGGGAGACAGCTTTAGCTCAAGTATTA 63
QY 689 TCAGCAGTGTATATCATAGTAACTTAATGGGAGCGTATCATAGAAATATTAACATATA 748
DB 64 GCCGAAATGACACAGAGTCTGTATTTGGGAAATTAACAGAGCAATGATGAATA 123
QY 749 AACCTTACGAAAAAATTTAAATTTTCCCATATATCATCTACAGCAACATATAT 808
DB 124 AATACAAACACCCACAAACATTAATTTACAGAAATTAATCTTACTTCAATTCAG 183
QY 809 GCAACCAAGCAATATAGAGAAATATCTGACAAACAATAC-ATGATTAATCTTAACTA 866
DB 184 GCAACCACTGATTTAGCACTGCTTTAGTATGATGTGATGGGTAAGTGTGTGATGCA 243
QY 867 TTCCACACAACTATAGCACCATATGTACAAATTAACAAACAGATATGTGA 926
DB 244 ACTAAAGCAATTCGTACGTAGCAAGAAATGGCCCAACCTTTAAACAAACCAATT 303
QY 927 AAATATCCCAATTTAATTTGTAGTAAAGTATGAAATTAATCACTCAATTTTCCA 986
DB 304 ATTATTTATGCAAGTAAAGGCTCGAACAGGTACACATTAAGGATTTGAAATGATA 363
QY 987 GTGAATTCGCAAGAAATTTTAATTAATCAATTTTATATCTCTGTGCTCAAGTT 1046
DB 364 GCGGAAGAGATTCAGCAGA--AAACGGCAAGGATGTGTTTATCTGTGTCAGGCC 421
QY 1047 TTGCTAAAGAAATTCAGAACATCTTCTGTATATAGTACTGCTGATATATAAG 1106
DB 422 ATGCAAGAAAGTGGCTGTATATGACATTAACACATTAAGCAAGCAAAATTTAG 481
QY 1107 AACTGTGTAATCAATGATAGAAACAATAGTAATGATGTTCTTAATAATTAATCAATC 1166
DB 482 CAGACGCTGTATGTAGAGAGAGTGTTAAGATGATTTCTTGAATTTATTAAGATG 541
QY 1167 AAGATATTATGCTGACAGATTGAGCTGATTAAGACATATATGCAATTCATGTG 1226
DB 542 ATGATGTATGTTGTGACAAACAGGGCGGCTTAATAAATATTATTCCTTGGCGCGCG 601
QY 1227 GAATTAATGCTGAAAAAATTAGTATAATATGCTGTCTACTGTTATTAATTAAGCA 1286
DB 602 GAGCGATTTCATGGCTTGGCTTGGTGTATATGCAAGCGCAATTATGACTCGGTGTT 661
QY 1287 TGAATGAATTAACACTATATATAGCAAAAAATCTTCAATAGATCTTCAATCTTAA 1346
DB 662 TAGCAAAATTTAGTCGTTTGGGCTTGCATGGG-----AGCAAACTTAACTTAA 715
QY 1347 TTGGTTCATCATGCTTGGAGATCTAATTAATTAACATGACAGAACTTACAGCAATA 1406
DB 716 TTGCTTTAAGTGTGTTGGGGAATTAATGTTTACGTTGACAGATGTCATTCGAAAT 775
QY 1407 TTGCTTTGACACTGAATATGAAAAAGTAAATATTAATATACATTAATATGATCAAC 1466
DB 776 GGGCTGGCGGTAATTTATTTAGAAAAAGCCATTAAGTAAAGTCTTGAAGAAACATGG 835
QY 1467 TAAAGCTTGTGAAGAACAGTACTGTAAACCACTGATATCATTAGCAAAAAAATTA 1526
DB 836 GCATGATTTGGAGAGTGTTCACAACTAAAGAGCTATGAATTTGGCAACAACTAG 895
QY 1527 ATGTAGAACTACCAATTTGCAATCTTATTTCAATTTTATTAATGAGATATATCACTAG 1586
DB 896 AAGTGAATGCCATACAGAAACAATCTAACATGCTTGTACAAATGAGAAATGTAC 955
QY 1587 ATAAAGCATATCAACATATTAAT 1610

DB 956 AACAGCAGCAAAAGAAATCATGT 979

RESULT 23
US-09-815-242-3869
Sequence 3869, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:

APPLICANT: Heselbeck, Robert
APPLICANT: Olsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT FILING DATE: 2001-03-21
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3869
LENGTH: 1026
TYPE: DNA
ORGANISM: Enterococcus faecalis
US-09-815-242-3869

Query Match 3.1%; Score 147.2; DB 9; Length 1026;
Best Local Similarity 49.1%; Pred. No. 6,6e-14;

Matches 480; Conservative 0; Mismatches 488; Indels 10; Gaps 3;

DB 635 AAAATCAGTATTTAGGTGCGAGATCTTTGCGACAGCAATGCAATTGCACTGTGACCA 694
4 AAAGTCGCTGTTTGTGCTCAGGTTCTGCGGGAAGCTTTAGTCAAGATTAGCGGAA 63
QY 695 CATGTATATCATCTAATCTATGGGAGCTGATCATATAAATTTACCATTTAAACT 754
DB 64 AATGACACGAGGCTGTATTTGGGAAATAAACAAGCAAAATGATGAATATAACA 123
QY 755 TACCGAAAAAATTTAAATTTATTTACCATATCATCTACACAGCAACATATGCAAC 814
DB 124 AAACACACCAAAACATTTATTTACAGATTTAATCTTACTTCAATTCAGGCAACC 183
QY 815 AGCAATATAGAGAGATATCTGACAAACATAC-ATGATTTATCTTAATTTCTT 872
DB 184 ACTGATTTAGGCACTGCTTTAGTATGATGTGATGGGAGTACTGTTGGTACCACTAAA 243
QY 873 CACACAAATTTACGACCATATGTCACAAATATACAAACAAACAGCATATGTATAATA 932
DB 244 GCCATTGCTGATGACACAAAGAGTGGCCCAACCTTTAAAAAACAACATATTTAT 303
QY 933 CTCCAATTTAATTTGTATGTAAGGTATCGAAATTTACATCAATTTCCCGTAAA 992
DB 304 CATGCAAGTAAAGGCTTGAAACAGGTACACATTAAGCGATTTGAGAAGTATGACGAA 363
QY 993 TAGCAGAGAAATTTACATATTAATCAATTTTATCTGTGTCAGATTTTGCTA 1052

DB 364 GAGATTCAGACAA--AAAACGCAAGGATTTGTTTATCTGTCCAGCAAGCAG 421

QY 1053 AAGAAATTCAGACAACTTCTCTGTATATAGTACTCTGTGTATTAAGAAGCTG 1112

DB 422 AAGAGTGGCTGTATGACATTTACACACATTTACAGCAAGCAAGAAATTTAGCAGAG 481

QY 1113 GTGAATCATTTAGTAAACATTAAGTATGATGTTCTTAAAAATATATACCATCAAGATA 1172

DB 482 CTGTATATGTCAGAGGTTGTTTATGATGATTAATTAATTAATTAATTAATTAATTAAT 541

QY 1173 TTATAGGTGACAGATTTGAGCTGATTAAGCACTAATTTGATTTGATTTGATTTGATTT 1232

DB 542 TGATTTGTGTAAGAAACAGGGGCGCTTTAAAAATTTATTTGTTGGGCGCGGAGCA 601

QY 1233 TCGCTGAAAAAATTTAGTATATATGCTGTTGCTACTGTTATTAATTAAGCATGAATG 1292

DB 602 TTGATGCTTTAGGCTTTGATTAATGCAAGCGCAATTAATGATCTGTTGTTAGCAG 661

QY 1293 AAATTAACACTATATATATAGCAAAATCATTCATTAATGATCTTCAATTAATTTGCTC 1352

DB 662 AAATTTAGCTTTAGGGGTTGCGATGG-----AGCAACCATTAATTAATTTATTTGTT 715

QY 1353 CATCATGCTTTGAGATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1412

DB 716 TAAAGTGTGTTGGGATTTAATTTATCTGTACAGTGTCAATTTCCGAAATTTGCGCTG 775

QY 1413 TTGAGCTAGAAATAGGAAAGGTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1472

DB 776 CGGTATTTATTTAGGAAAGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 835

QY 1473 TTGTGAAGAACCACTACTGTAAACCACTGATCATTAAGCAAAAAAATTAATGATG 1532

DB 836 TTGTGAAGAGTGTTCACAACTAAAGCAGCTGATTAATTTGCAACAACTAGAGTGC 895

QY 1533 AACTACCAATTTGCAATATTTTAAATTTTAAATTAATTAATTAATTAATTAATTAATTA 1592

DB 896 AATGCAATCAGCAAAACATTTCAATGCTTTGATCATGACGAAATGATGATGATGATG 955

QY 1593 CCATATCAAAACATTTAT 1610

DB 956 CAGCAAAAGAAATCATGT 973

RESULT 24
US-10-311-455-836/C

Sequence 836, Application US/10311455
Publication No. US20030143606A1

GENERAL INFORMATION:

APPLICANT: OLEK, Alexander

APPLICANT: PIEPENBROCK, Christian

APPLICANT: BERLIN, Kurt

TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detect

FILE REFERENCE: 5013.1014

CURRENT APPLICATION NUMBER: US/10/311,455

PRIOR FILING DATE: 2002-12-16

PRIOR APPLICATION NUMBER: PCT/EP01/07537

PRIOR FILING DATE: 2001-07-02

PRIOR APPLICATION NUMBER: DE 10032529.7

PRIOR FILING DATE: 2000-06-30

PRIOR APPLICATION NUMBER: DE 10043826.1

PRIOR FILING DATE: 2000-09-01

NUMBER OF SEQ ID NOS: 2424

SEQ ID NO 836

LENGTH: 6132

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-311-455-836

Query Match 3.1%; Score 147.2; DB 14; Length 6132;

Best Local Similarity 44.2%; Pred. No. 1.3e-13;
Matches 1069; Conservative 0; Mismatches 118; Indels 31; Gaps 10;

```
QY 281 AGAAGACAGCTGATTCAGAAAGAAATACCTCTCTCTGTCATCCATATTTTCCATTTAT 340
Db 5037 AACCAAAAACCTTTATTTCTTCACACCTCTTAATTTTCTCCCAAAATCCCTAAAAAAT 4978
QY 341 TTACAAATGATAGAGTCTTTATTCGACATACATATTAATTCATCAACAAACAAATTA 400
Db 4977 ATAAAAATATATAA---TTTATTAACACTTATATTAATTAATTAATTAATTAATTA 4921
QY 401 TAGACATATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 460
Db 4920 AAAACATTTACTACCACTTAAAAAATTAATAAATTAATAAATTAATAAATTAATTAAT 4861
QY 461 TGAACACTTAATCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 520
Db 4860 TTATACACC--ACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4802
QY 521 AATTCATCTGATGATGATTAACATATGATACATTTCTTAACCTGATATATTAAGA 580
Db 4801 ACTACCAAAACAACTCAAAAAAATTAATAAATTAATAAATTAATAAATTAATAAATTA 4742
QY 581 ATAAATATATATATATTTTGAATGCTTCTGATATTAATTAATTAATTAATTAATTA 640
Db 4741 CTATTCATATTTAC--TAAATCTCTAACAAACAAATTAATTAATTAATTAATTAATTA 4684
QY 641 AGATTTTAGTGAGATCATTTGGACAGCAATGCAATGCACTGTCAGCATGCT 700
Db 4683 AAAATATCCATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4624
QY 701 ATATCACTTACTATGAGGAGCGATGATGATGATGATGATGATGATGATGATGATGAT 760
Db 4623 ATACAAAAAAATCCCAATATCCACACAAATTAATTAATTAATTAATTAATTAATTAAT 4564
QY 761 AAAATTTAAATATTTTACCCATATCATCTACAGACAAATATATGCAATATGCAAC 820
Db 4563 CAATTTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4504
QY 821 ATGAGAGATTTATTTGACAAATATGATGATGATGATGATGATGATGATGATGATGAT 880
Db 4503 AAAAAAAACCTAAA---AAACATTTCTTTATTAATTAATTAATTAATTAATTAATTA 4448
QY 881 TTAGCACATATGATGACAAATTAACACAAACAGCATATGATTAATTAATTAATTAAT 940
Db 4447 ATAAAACTAATTTAACAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4388
QY 941 TTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1000
Db 4387 CTATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4328
QY 1001 GAAATTTTACATATATCCAAATTTTATCTCTGTCGACAGTTTGTCTAAAGAAAT 1060
Db 4327 CAATTTCTCCCAAAACATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4278
QY 1061 GCGAACAATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1120
Db 4277 TCAATTAATCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4218
QY 1121 TTGATAGAAACAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1180
Db 4217 AAATCCAAAAAACAACCAATTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 4158
QY 1181 GTACAGTGTGAGTGTGATTAAGAAATATATGATGATGATGATGATGATGATGATGAT 1239
Db 4157 TCTCAATTTCAAACTTACAACTACATTAATTAATTAATTAATTAATTAATTAATTAAT 4098
QY 1240 AAAAAATTTAGTAAATATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1299
Db 4097 AAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4038
QY 1300 AACCTATATATGCAAAAAATCATTCATATGATCTTCACTATTAATGCTGCTGCTGCT 1359
```

```
Db 4037 TTTAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3978
QY 1360 TCTTGAGATCATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1419
Db 3977 TCTTCACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3918
QY 1420 AGAAATAGAAAAAGTGAATATATA--TACATTAATTAATTAATTAATTAATTAATTA 1477
Db 3917 CTACCTATATACATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3858
QY 1478 GAAAGAACAGTACTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1537
Db 3857 AATCCATTAATCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3798
QY 1538 CCAATTTGATATCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1597
Db 3797 AATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3738
QY 1598 TCAACATATTAATCTTATGCTATGATGATGATGATGATGATGATGATGATGATGATG 1657
Db 3737 ACATTAATTAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3678
QY 1658 GTCAAAAAATATAGATTAACAGTTTACACACATATTAATTAATTAATTAATTAATTA 1717
Db 3677 TACAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3618
QY 1718 TAACTGCACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1777
Db 3617 ATATTAACAACTCTTCCAACTGCAACAAATTAATTAATTAATTAATTAATTAATTA 3558
QY 1778 CTACCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1837
Db 3557 CAAAAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3498
QY 1838 TATTAATTAATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1897
Db 3497 AATATATCTCAACTCTATCTCTAATTAATTAATTAATTAATTAATTAATTAATTA 3438
QY 1898 CACTACATCATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1957
Db 3437 CTTTACACCCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3378
QY 1958 TACATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2017
Db 3377 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3318
QY 2018 ATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2075
Db 3317 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3258
QY 2076 TTATGCCCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2135
Db 3257 CCAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3198
QY 2136 CTGCAACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2195
Db 3197 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3138
QY 2196 ACCACCTACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2255
Db 3137 ACCCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3078
QY 2256 GCATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2315
Db 3077 TTATTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3018
QY 2316 ACTTCAACAACTCAAAAGTTCTATTAATTAATTAATTAATTAATTAATTAATTAATTA 2375
Db 3017 AACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2299
QY 2376 CAACATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2435
Db 2298 TATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2899
```


Best Local Similarity 49.8%; Pred. No. 2.1e-12;
Matches 484; Conservative 0; Mismatches 473; Indels 14; Gaps 4;

```
QY 613 GATATTAATAGTGTAAAAAATGTAATCAATATTTAGGTGAGATCATTTGGACAGC 672
Db 2065487 GAAAAATATATGACACAGAAAAAGTGTCTATTTCTGGCGCTGGAAGTTGGGAACAG 2065428
QY 673 AATAGCAATTCGCTGCGACATGATATGATTAAGTAACTTAATGGGCGGATCATAG 732
Db 2065427 ACTGCGCTTCTCTGCTGATATATATATTAACCAAGTATTTGGGAACTTATGATTA 2065368
QY 733 AATATTTACATATAAACAATTACGAAAAATTTAAATATTTTACCACATATCATCT 792
Db 2065367 AATGTGATGAAATTAATACGAATGCAACGAATAGTCACTATTTGCCAGATATATTTT 2065308
QY 793 ACCAGACATATATATGCAACCGACATATAGCGAATATTTCTGACACAAATATCATG 852
Db 2065307 ACCAATGAGGTAAAGCGCATTTATCACTGTATGAGTATATAGATGTGCTGAAT--- 2065251
QY 853 TATTAATCTTAATCTTCTACACACATTTACGACACATATGATACAAATACACACAA 912
Db 2065250 TGTGCGATGTATTTCAACAAATGCAATGCGTGTGTTGTAGAGACCTAAATGAGC 2065191
QY 913 ACAGCATATGTAAAAATCTCCATATTTATTTAGTAAAGGATGCAATTTACATC 972
Db 2065190 GCTGAA-AGAACCACTATTTTATGTCATGTATGTAAGGAGATGGAACCAACAAATC 2065132
QY 973 ACTCAATTTCCAG----TGAATAGCAGAAAGAAATTTTACATATTAATCCATTTT 1028
Db 2065131 TTGAAATGTGAAAGTATTTGAAGAAAGATGATGCTACAAACCTAAAGCTCTGTTG 2065072
QY 1029 TACTCTGTGCTCAAGTTTCTCTAAAGAAATGCGAAGATCTTCTCTGTATGATGAC 1088
Db 2065071 TTCTTTTCAGGCGCTATCTACGCTGAAGAAAGTACTCTTGTGATCAACAAACGCTTTG 2065012
QY 1089 TTGCTGTGATATTAAGAACTTGTGATATCATTTGATAGAAACAATAGTATGATGTC 1148
Db 2065011 CTGAGCTGTAAAGTTTGAAGACGCAAAATGCTGCAAAATGCTTTATTAATTAAT 2064952
QY 1149 TAAAAATATATATACATCAAGATATTTATAGTGTACAGATGAGCTGATTAAGAA 1208
Db 2064951 TACGATTTTATAGAAATGATGATGTCATGCTGCTGAATTTGAGAGCGCTTAATA 2064892
QY 1209 TATTTGCAATGATGATGATATATGCTGGAAGAAATTTAGGTATATATGCTGTGTA 1268
Db 2064891 TTTTCGCACTTGGCGAGATCTTGTATGCTGATGCTGATACGCGATATATGTAAGCG 2064832
QY 1269 CTGTATTAATAGGATGATGAAATTAACACTATATATAGCAAAAAATCATGCA 1328
Db 2064831 CACTATATGACGCGGATGCTGAAATTAATCTGCTGAGATGTGCTGCGCTCTA--- 2064775
QY 1329 TAAATCTTCAATATTAATTTGCTCATGCTGCTGAGATCATTAATTAATTAATTA 1388
Db 2064774 ---ATCACAAACCTTTTACGGGCTAACGTAATGATGTAATCTTAATCTTGACAA 2064718
QY 1389 CAGACATTCACCAATATGCTTTTGAAGTAAATAGAAAAAGGTAGAAATATTAATA 1448
Db 2064717 GTGTGATTCAGCTAATTTGGCGGTGGAACATGCTGCTGAAGCGCAATTTAGATG 2064658
QY 1449 CATTAATATATCAACCTTAAGCTTTGGAAGAAACGATCTGTAACCACTGATAT 1508
Db 2064657 AAGCTTTTGAAGAAATGAGTATGCTGTAAGAGTGTGCTGACGCAAGAGAGTTCA 2064598
QY 1509 CATTAACAAAAAATTTAATGTAAGTACCAATTTGCAATTTTATTAATTTATAC 1568
Db 2064597 GTTGGCGAAAAAATATGATATGATGCAATTTACCAATGATTTTAACGGAATTTAT 2064538
QY 1569 ATGAGATATA 1579
Db 2064537 TCGAGATATAA 2064527
```

RESULT 27

```
US-10-240-453-206/C
; Sequence 206; Application US/10240453
; Publication No. US20030148326A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIENENROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA
; TITLE OF INVENTION: Transcription
; TITLE OF INVENTION: by Means of Assessing the Methylation Status of Genes Associate
; FILE REFERENCE: 5013.1003
; CURRENT APPLICATION NUMBER: US/10/240.453
; PRIOR APPLICATION NUMBER: 2002-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/03973
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 350
; SEQ ID NO 206
; LENGTH: 11745
; TYPE: DNA
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (9105)
; US-10-240-453-206

Query Match 3.1%; Score 145.4; DB 14; Length 11745;
Best Local Similarity 43.8%; Pred. No. 3.3e-13;
Matches 968; Conservative 0; Mismatches 1241; Indels 18; Gaps 6;

QY 328 ATTTCCAAATTTATTAATATGATAGAGTTCTTTATTCGACATATATTAATCAAT 387
Db 6029 ATTAAATATATCTAACCAATATATATATCTTTAATATATTAATTAATTAATTAATTA 5970
QY 388 CAAACAAATATATAGTACTATCTATTTATTAATTAATTAATTAATTAATTAATTAATTA 447
Db 5969 AATTAATTAATTAATCTCTATTAATTTATTAATTAATTAATTAATTAATTAATTAATTA 5910
QY 448 ACTCAATATTAATGACACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 507
Db 5909 AAATTAATCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5850
QY 508 TCAATTTAGCATTAATTCATCTGACATGATTAACATATGTAACATTTCTTAACAC 567
Db 5849 AATTAATTAATTAATCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5790
QY 568 TAGATTAATTAATTAATTAATTAATTAATTAATTTTGAAGTGTGCTTTGATATTAATTA 627
Db 5789 AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5730
QY 628 AAAAAATTAATCAATTTTATGAGTGAAGATCATTTGGCAGCAATGCAATTTGCACT 687
Db 5729 AAACAAATTAATCAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5670
QY 688 GTCAGCATATGATATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 747
Db 5669 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5610
QY 748 AAACACTTACCGAAAAATTTAATTAATTTTATTAATTAATTAATTAATTAATTAATTA 807
Db 5609 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5550
QY 808 TGACACAGCAATATAGCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 867
```


| | | | |
|----|------|----------------------------------------------------------------|------|
| Db | 5549 | ATAAAAATCATAAAAATATCATTAATATATTAACCAATTAATTTATTCGAAACACTTAATTA | 5496 |
| Qy | 868 | TCCTACACAACATTAAGCCACCATATGATACCAATTAACAACGAACGACATATGTGTA | 927 |
| Db | 5489 | ATCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT | 5430 |
| Qy | 928 | AAATACCTCAATATTTAATTTGTAGTAAAGTATCGAAATTAACATCACTCAATTTCCGAG | 987 |
| Db | 5429 | AACCTTAACCAATTAATCAATTAATCAATTAATTAATTAATTAATTAATTAATTAATTAAC | 5370 |
| Qy | 988 | TGAATATGAGAGAAATTTTATCAATATATACCAATTTTATACCTCTGTGCGCAAGTT | 104 |
| Db | 5369 | TAAATCATCAACAAAAAACACCAATACAAAATCTAAAAAAAATTTCAAAAAAAA | 5310 |
| Qy | 1048 | TGCTAAGAAATTCGAGAACATCTTCCCTGTAGTATGATCTGTGTGTGATTAATAAAGA | 1107 |
| Db | 5309 | TAAACATCACTACAAAAACCTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT | 5250 |
| Qy | 1108 | ACTTGCTAAATCATGATGATGAAGAATTAAGTATGATCTGTAAATTAATTAATTAATTAAT | 1167 |
| Db | 5249 | AAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA | 5190 |
| Qy | 1168 | AGATATTAATGAGTGCAGATTTGAGCTGCATTAAGAACAAT--AATTGCAATTCGATGT | 1225 |
| Db | 5189 | ATAAAAAAATATCAAAATATATTAATTTTAACTATATTAATTAATTAATTAATTTATTT | 5130 |
| Qy | 1226 | GGAATTAATCGCTGGAATAATTTAGTATATATGCTGTGCTAGCTGTATTAATCAAGAC | 1285 |
| Db | 5129 | CTAAATACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA | 5070 |
| Qy | 1286 | ATGAATGAATTAATAACAATAATATATAGCAAAAAATCAATCAATAGATCTTCATACATTA | 1345 |
| Db | 5069 | TTAAAAAATCAACCTAATCACTAATTAATTAATTAATTAATTAATTAATTAATTAATTA | 5010 |
| Qy | 1346 | ATTGCTCATCATGCTTGAGATCTAATAATTAACATGACAGCAACATTCACGCAT | 1405 |
| Db | 5009 | AATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA | 4951 |
| Qy | 1406 | ATGCTTTTGGACTAGAAATAGGAAAGGTAGAAATATTAATACATTAATAGATCAAC | 1465 |
| Db | 4950 | -----TTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA | 4897 |
| Qy | 1466 | CTAAAGCTGTGGAAGAACACGATCTGTAAACACGATATCATCTTGCAAAAAAATTT | 1525 |
| Db | 4886 | ATTAACATTAATACTTAATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA | 4837 |
| Qy | 1526 | AATGTAGACTACCAATTTGCATATCTATTACATTTTATTAACATGAGAAATATATCACTA | 1585 |
| Db | 4836 | AATTACACCTA--AAGTTTAACCTAAAAAAACGAAATTAATTAATTTCTATTTCTAA | 4778 |
| Qy | 1586 | GATAAGCCATTCACCAATTAATCTTAAGCTATCATCTTTGATATTTCAACACAT | 1645 |
| Db | 4777 | TAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT | 4718 |
| Qy | 1646 | GATAAATTAACGGTCAAAAAATATGAAATACAAAGTTCAACAATATATTAATAAGACAGTCA | 1705 |
| Db | 4717 | ATCAATTAATCTATTTTAACCGTATCAAAAAATCATATACATCTAACCTTAATAATATTA | 4658 |
| Qy | 1706 | AATACCCATCGTTAAACATCCCACTTAAGATGTAACATTTCAATTAACAACTTAAGTATA | 1765 |
| Db | 4667 | AATACCACTATATTAATTAATCTTA--AAGTTAAAAAAAATTAATTAATTAATTAATTAAT | 4600 |
| Qy | 1766 | ACTTAAATTTCTAGCTAATTAATTAACAAATTCCTATCTCATTTAACCA | 1825 |
| Db | 4599 | TTTAAAAATACATCAATTAATTAATTTTAAAAACCTTAATAATTAATCAACACTCA | 4540 |
| Qy | 1826 | TTGTATCAGATTAATTAATTAAGTCAAAATTAAGCTTTCTTAATTTAATTAATTAATTT | 1885 |
| Db | 4539 | AAAAAAAAAACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA | 4485 |
| Qy | 1886 | TGCACACTTTTCACTACATCATGTATTAATTAAGAACATTAATTTTACATTAATAATTAAT | 1945 |

| | | | |
|----|------|----------------------------------------------------------------|------|
| Dc | 4484 | TCCTACATTTAAAAATCTAATTAATAAAAAACAAAAAACAAAAAATAAAAAATACCAT | 4445 |
| Qy | 1946 | AAAAAGAACCAATGCATATATTTAAATTAAGTAAACACCTTAATTTCTGCGTAATTTGCAT | 2005 |
| Dc | 4424 | TAAATATAAAAAATAAATCTATACAAATTAATATTCAAAAATACATATAAAAAAACCTTC | 4355 |
| Qy | 2006 | AGATGTCCTGCATATATAATTCTATCATCTTGACATTTTCAACATATAAATATGATA | 2065 |
| Dc | 4364 | AAAACAATAAAAATAAACCAATTACATATATACCTTTAAATCCAAAAACAAAACTAAA | 4305 |
| Qy | 2066 | AACGTGTTTTTAAGCCCATTAAGTCTTAATTAATAATGCTAATTAACCCAAAAAGTAAA | 2125 |
| Dc | 4304 | ATCATTTAATTTTATCAATATATATACGCATTAATATTTAATTAATTAATAAACCTTTTA | 4245 |
| Qy | 2126 | ATACACATTCCTGCACATATATGAAGAAGTACATACAGCTCAATTAATTTAAACAAT | 2185 |
| Dc | 4244 | TTAAATATTAATAAAAAATATATATCTTATTTAAATATAATTAATAAAAAATTAATTAAT | 4185 |
| Qy | 2186 | TTCTTAATATACACACCTACTCTAAACATAGAAACCAATAGT-ATATCATACCTAAGTAG | 2244 |
| Dc | 4184 | AAAAACAACAATATTAACAACAATTTAAAAAATCTACTAATTAATAAAAAACAAAAA | 4125 |
| Qy | 2245 | ATATCCCTACAGCATATATCAAAATTAATAAGAAATAACGTATGATATTTATCAATTA | 2304 |
| Dc | 4124 | TAAAGCTTATATTAATAATATTAATTCAAAAAATTTCTTTAAATATATATTTAAAA | 4065 |
| Qy | 2305 | CAAAATTTCCAACTTCAACAACAACAACCTCAAGTCTATATATAACAATTCAGAATTAAT | 2364 |
| Dc | 4064 | CAACTCATATATATCAATATAAAAAATTAACCCAAATAAAAAACAACAATATTAACAACA | 4005 |
| Qy | 2365 | ATATACTTCACCAACATATGTTATCAAAATTCCTATATAACATCATATCCTTTTAAAAAG | 2424 |
| Dc | 4004 | AAAAAAAATCTTACAAAAATTAATAATCTTTTAAAAAACAATTAAAAATTTAAAAAC | 3945 |
| Qy | 2425 | ATGACTTCACATTCATACATAAATAATTAATTAACATGCAATGTTTATTCAGAAACATTACA | 2484 |
| Dc | 3944 | ACAAATAAAAAACTAAACCTTAATAATAAAAAAACAACCTTCATTTTAACAAAAAAAA | 3885 |
| Qy | 2485 | CAATCCAGAATTTTCATTTAGCGTATATGCGAAGCTTCCATTCACAACAATTCAGC | 2544 |
| Dc | 3884 | AAATATTTTATATTAATTAATTTAAAAACAATTTACTAATAATTAATTAACAACAAT | 3825 |
| Qy | 2545 | AGTTTCGT 2551 | |
| Dc | 3824 | AATTTCTT 3818 | |

RESULT 28

US-10-311-455-986/c
Sequence 886 Annotation MS/10311455

; Publication No. US20030143606A1

GENERAL INFORMATION:

APPLICANT: OLEK, Alexander
PIEPENBROCK, Christian

APPLICANT: BERLIN, Kurt

TITLE OF INVENTION: Diagnosis of Diseases

FILE REFERENCE: 5013-1014

CURRENT APPLICATION NUMBER: US/10/311,455

CURRENT FILING DATE: 2002-12-16

PRIOR FILING DATE: 2001-07-02

PRIOR APPLICATION NUMBER: DE 10032529.7

PRIOR FILING DATE: 2000-06-30

PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: DE 10043826.1

NUMBER OF SEQ ID NOS: 2424

SEQ ID NO 986
LENGTH: 7567

```

1 LENGTH: 755
2
3 TYPE: DNA
4

```

ORGANISM: Artificial Sequence

OTHER INFORMATION: chemically treated cre

SECRET
NO FOREIGN DISSEM

US-10-311-455-986

Query Match 3.0%; Score 144.2; DB 14; Length 7597;
Best Local Similarity 47.4%; Pred. No. 4.3e-13;
Matches 556; Conservative 0; Mismatches 613; Indels 4; Gaps 4;

1294 AATTAAACGCTATATGCAAAAATATCATAGTCTTCTATATTAATGGGCC 1353
1272 AAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1213
1354 ATCATGCTTGGAGATCTAATTTAAACATGTAACAAGACATTCACCAATGGCTT 1413
1212 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1153
1414 TGGACTAGAAATGAGAAAGTAAGAAATTAATTAATTAATTAATTAATTAAT 1473
1152 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1093
1474 TGTGAAGAACACGACTGTAATAACCATGATATCATAGCAAAAATTAATTAAT 1533
1092 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1013
1534 ACTACCAATTTGATATCTATTTTACATTTTATTAATTAATTAATTAATTAAT 1593
1032 AATACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 973
1594 CATATCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1653
972 CATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 913
1654 AACGTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1713
912 AATACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 853
1714 TGTGATCACTCACTAGTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1773
852 AATACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 793
1774 TTTCCTAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1832
792 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 733
1833 CGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1892
732 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 673
1893 TTTCCTAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1952
672 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 613
1953 ACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2012
612 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 553
2013 CTGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2071
552 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 493
2072 TTTCCTAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2131
492 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 433
2132 ATTCCTGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2191
432 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 373
2192 TATTAACCACTACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2250
372 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 313
2251 CTGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2310
312 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 253

QY 2311 TTCCAACTTCACACAAAACTCAAGTCTATATTAACAAATTCAGAAATTAATATAC 2370
DB 252 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 194
QY 2371 TTCACCAACAGATAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2430
DB 193 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 134
QY 2431 TCACAAATTCATACATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2463
DB 133 AAAAAATACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 101

RESULT 29

US-10-312-841-2/c
; Sequence 2, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: EpiGenomics AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MR
; FILE REFERENCE: E01/1208/KO
; CURRENT APPLICATION NUMBER: US/10/312, 841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 2
; LENGTH: 3673778
; TYPE: DNA
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: chemically created genomic DNA (Homo sapiens)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (379615)
; US-10-312-841-2

Query Match 3.0%; Score 141.6; DB 14; Length 3673778;
Best Local Similarity 43.8%; Pred. No. 1.2e-11;
Matches 1001; Conservative 0; Mismatches 1269; Indels 18; Gaps 8;

QY 374 ATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 433
DB 2092232 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2092273
QY 434 AGTCCTTACGACAACTCAATTAATTAATTAATTAATTAATTAATTAATTAAT 493
DB 2092272 CCACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2092213
QY 494 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 553
DB 2092212 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2092155
QY 554 ACATTTCTTAACCTAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 613
DB 2092154 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2092095
QY 614 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 673
DB 2092094 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2092035
QY 674 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 733
DB 2092034 TTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2091975
QY 734 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 793
DB 2091974 ACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2091920
QY 794 CCAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 853
DB 2091819 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2091860
QY 854 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 913

| | | | |
|----|---------|----------------------------------------------------------------|----------|
| Db | 2091859 | AAATCCAACTTACAAATAACATTTAAAAACAATAAAAATCAAAAAAATTTAAAAACCCAAAC | 20918000 |
| Qy | 914 | CAGCATATGTGTAATAAATATCCCAATATTAATTTGTACTAAAGATCGAATATACATCA | 973 |
| Db | 2091799 | AATACTCTTCACTATATATATTAATTAATTTATTAACCAATACATATATATATCAATTA | 20917400 |
| Qy | 974 | CT--CAAAATTTCCAGTAATATAGCAGAGAAATTTTTCACAATATTAATCCAAATTTTATA | 1030 |
| Db | 2091739 | AAAAAAAATATATCTTTTCATCAACAATATCAAAAAACAACAATATCCCTTATACAAA | 20916800 |
| Qy | 1031 | CTCTCGTCCAAAGTTTGGTAAAGAAATGCGAGAACTCTTCCGTGAGATAGTACTT | 1090 |
| Db | 2091679 | AAATTAATCTAATATCCCTACTTACACCAATATTAATAATTAATCCCAATATATCAAAAC | 20916200 |
| Qy | 1091 | GCTGGTGATATTAAGAACTTGGTGAACTATTCATGTAGAAACAATATAGATATGTTCTA | 1150 |
| Db | 2091619 | CTAATATATAAACCTTAATATTAATAAAAACTTTAATAAAAAACATTAACATTAATCTTCAT | 20915600 |
| Qy | 1151 | AAAAATATATACATCAAGATATTTATAGTGTAACGATGGAGCTGCATTTAAAGACATA | 1210 |
| Db | 2091559 | AATCTTAATTTAATTAATAAAATCTTAATCTTAATATATACAAAAACAACAACAAAAA | 20915000 |
| Qy | 1211 | ATTGCAATTGCATGTGAAATATCGCTGGAAAAATTTAGGTAAATATGCTGTTGCTACT | 1270 |
| Db | 2091499 | AAAAATTAATAATTAATTTAAACATCATCAAAATTTAAAAATATATTAATCTAAAAACACA | 20914400 |
| Qy | 1271 | GTTATTAATTAAGGATGAATGAATTAACAATTAACATATATATAGCAAAAAATCATTCATA | 1330 |
| Db | 2091439 | TCAAAAAATTAATAAAAAATCTAATTAATTAATAAAAAAATTTTACAAATCTTAATCTTA | 20913800 |
| Qy | 1331 | GATCTTCATACATTAATTTGTCATCATCATGTCTGGAGATCTAATATTAATTAACATGTACA | 1390 |
| Db | 2091379 | CAAAAAAAATCTATATCTAATAATATATAAAAATTAATTAATCTCAATATTAATTAAT | 20913200 |
| Qy | 1391 | GAACTTATAGGCAATATGCGTTTGGATAGTAATAATGSAAGATGAGAAATATTAATACA | 1450 |
| Db | 2091319 | TAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA | 20912600 |
| Qy | 1451 | TTAATAGATCAAACTTAAGCTTGGTAGAGAAACCACTACTGTAAAAACAATGATACA | 1510 |
| Db | 2091259 | AAACCATTAACACATTAATAAACATATCAATCATTTAAACATCAAAAAATTAACACATCA | 20912000 |
| Qy | 1511 | TTAGAAAAAAATCTTAATGTAGAACTACCAATTTGCAATCTATTTACAATTTATTCAT | 1570 |
| Db | 2091199 | ACCCAAATTAATTAATCTAT--TTCCACCCCACTAAATACCTAATTAATTTAAAAAATATA | 20911410 |
| Qy | 1571 | GAGATATATCACTAGATTAAGCATATCAACAATTTATCTTATGCTATCATCTTTGA | 1630 |
| Db | 2091140 | ATTATATATTAATAAATATTAATAATCTAATAATACACACTTTACTAATTAATAATTTA | 20910810 |
| Qy | 1631 | TAAATTCACACATGATTAATTAACGGTCAAAATTAATAGATTAACAAGTTCACACATA | 1690 |
| Db | 2091080 | AAAAACATTAACCCCTTTAAAAAACAACCTAATCAATTAATTTAATCAATTAATTT | 20910210 |
| Qy | 1691 | TATTAAGCAGTCAAAATACCATCGTTACATCCAACTAAGATGTAACATATTTACATA | 1750 |
| Db | 2091020 | ATTACATTAACCAACAATCCCTCTCTAATAA---CAATATACCAAAAAAATTAATAAC | 20909640 |
| Qy | 1751 | ACAACTTAAGTATTAATTTCTGACTAATTAATTAATAACAATTCCTATCTGCT | 1810 |
| Db | 2090963 | ATTAATTCACATTAATAACCTTAATACCATATTTATTAACACATTTAATTAATCAATCCAA | 20909040 |
| Qy | 1811 | ATCTCAATTAACCAATTTGTTATCGATTAATAATTAAGTCAATTAACCTTTACTTATTT | 1870 |
| Db | 2090903 | ATAAAAAATTAATTAATTCACAACATTAATTAATAATTAATTAATTAATTAATTAATTA | 20908440 |
| Qy | 1871 | ATAAATATATGTGTGGACACTTTTTCAGTAC--ATCATGATATATATGACCAATATTT | 1928 |
| Db | 2090843 | ATATATATTCACAGATTAATAAATATTAATAATTAATTAATTAATTAATTAATTAATTA | 20907840 |
| Qy | 1929 | TACATTAATTAATTAATTAATAAGCAACATATCATATTTAAATTAG--TAAAAACCTATTT | 1987 |

| | | | |
|-----------------------------------------------------------------|---------|----------------------------------------------------------------|---------|
| Db | 2090783 | TTAAAAATATATCCGATGTAATAAAAAAAACAAACACAAAAAACCATATTAATTAATTT | 2090724 |
| Qy | 1988 | TACACGCTAATATATACATAGTGCCTGACATAATAATTCTATCACTCTGCATATTTT | 2047 |
| Db | 2090723 | CATTACATTAATAATTCGAAATTAATAATCTATTAATTAATTAACATTAATTAATTTT | 2090664 |
| Qy | 2048 | CAAACTAATAATATGATTAACCTGTTTTTTTATGCCCATTAGTCTTAATAAATTTGCTAA | 2107 |
| Db | 2090663 | CTTAAAACTAAAAATTTTAAATATAAATTTTCTCAAACTATAAAAAAATAAAAAATCT | 2090604 |
| Qy | 2108 | TTACCCAAAAAGTTAAATATACATCTCCGCAACAATATGAGAGAGATACATATACAGT | 2167 |
| Db | 2090603 | ATATTTAATTTAATTAATTAATAAACACACATTAATAATTTCTTAATAACCACTTAATTAAT | 2090544 |
| Qy | 2168 | CAAAATAAATTAACAAATTTCTAATATACACACTACTCTAAACATAGAACCAATAGTA | 2227 |
| Db | 2090543 | ATTTTAATATATACGAATTTATTACTATTAAATTAATATCTCAAAATTCCTTAATAATAT | 2090484 |
| Qy | 2228 | TATATACCTATAGTAATATATCCCTACAGCATGATCAAAATTAATAAGAAAAATCGTAT | 2287 |
| Db | 2090483 | TAAATTCATTTAATAACTTTAATATCATTTACACTCAACAAAACTACCAATTAATAACAC | 2090424 |
| Qy | 2288 | GTATATTTATACATTAACAAATTTCCAACTTCACACACAAAACTCAAGTTCATTATA | 2347 |
| Db | 2090423 | TCACATTAATTTACGTACGAATTAATAAAAAATTAACATACAAAAAATTAACCTAATACATA | 2090364 |
| Qy | 2348 | ACAAATTCAGAAATTAATATATCTCTCCACACATCTGTATCAATCTAATATACATC | 2407 |
| Db | 2090363 | ATATATCACCCTTAATTAATTAATTTAAATTTAAATTAATAAAAAAATTTAAATAATATTTT | 2090304 |
| Qy | 2408 | ATATCTCTTTTAAAGAGTACTTCACAAATTCATACATAAA-TAATTAATACATGCAATGT | 2466 |
| Db | 2090303 | ATACATATTTTAAATTAATTTATTTCTCTACTATTAACCAAAATTAATTTTACACTAAATTT | 2090244 |
| Qy | 2467 | TTATTCAGAAACATTAACAAATCCAAAGATTTTCAATTTACGGGTAAATCGTAGCTTC | 2526 |
| Db | 2090243 | TAAAAATTAATAATTTCAAAATCTTCAAACTCTACATCAATTAACATTAATTAACAAT | 2090184 |
| Qy | 2527 | CATTCCAACAATTCAGACAGTTCGTATACATTACCTCCAAAACGAGATACAGAGTTT | 2586 |
| Db | 2090183 | AATATTAAGTATCTAATAAATATCTCTCTAATAAAAACTAATCAATAACATACATACAAAA | 2090124 |
| Qy | 2587 | TAAATACGTCTTCAAAATCTTCACAGAGCTTACGTAATGTATACAGAAATACTTAGC | 2646 |
| Db | 2090123 | AAAAACAATACATTAATAAATTCAAAACAAACCTACATTAATTAATAAAAAAATAATTA | 2090064 |
| Qy | 2647 | ACTTAAAA 2654 | |
| Db | 2090063 | ACATAAAA 2090056 | |
| RESULT 30 | | | |
| US-10-398-221-3618 | | | |
| ; Sequence 3618, Application US/10398221 | | | |
| ; Publication No. US200400185141 | | | |
| ; GENERAL INFORMATION: | | | |
| ; APPLICANT: KUNST, Frederik | | | |
| ; APPLICANT: GLASER, Philippe | | | |
| ; TITLE OF INVENTION: Listeria innocua, genome and applications | | | |
| ; FILE REFERENCE: 344 702 - US | | | |
| ; CURRENT APPLICATION NUMBER: US/10/398,221 | | | |
| ; CURRENT FILING DATE: 2003-03-27 | | | |
| ; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061 | | | |
| ; PRIOR FILING DATE: 2001-10-04 | | | |
| ; PRIOR APPLICATION NUMBER: FR 00/12 697 | | | |
| ; PRIOR FILING DATE: 2000-10-04 | | | |
| ; NUMBER OF SEQ ID NOS: 4025 | | | |
| ; SOFTWARE: PatentIn version 3.0 | | | |
| ; SEQ ID NO 3618 | | | |
| ; LENGTH: 2791 | | | |
| ; TYPE: DNA | | | |
| ; ORGANISM: Listeria monocytogenes 4b | | | |
| ; FEATURE: | | | |

NAME/KEY: misc.feature
LOCATION: (1) (end)
OTHER INFORMATION: n can be any nucleotide: a,g,c or t/u
US-10-398-221-3618

Query Match
Best Local Similarity 49.4%; Pred. No. 1.5e-12;
Matches 480; Conservative 0; Mismatches 477; Indels 14; Gaps 4;

2.9%; Score 139.8; DB 15; Length 2791;
Best Local Similarity 49.4%; Pred. No. 1.5e-12;
Matches 480; Conservative 0; Mismatches 477; Indels 14; Gaps 4;

613 GATTAATTAATGTAATAAATGAAATTCAGTATTTAGTGCAGATCATTTGGCAGC 672
664 GAAAAATTAATATACACGAAAAAGAGTATCTTGGCGCTGGAGTGGGAAACAGG 723
673 AATGACATGACGTGACGACATGATATCATGTTAATGAGGAGCGATGATG 732
724 TCTTGACCTGCTGCTAGCTATATATCATCAACAGTATTTGGGAAACTTACATTA 783
733 AATATTAACATTAATAACCTTACGAAAAATTTAAATATTTCCACATATCATCT 792
784 AATGTAATGAATTAACGAAATCGACACGAAATAGCTATTTGCCAGATATATTTT 843
793 ACCAGAACATATATGCAACGACAAATATAGACAGTATTTCTGACAAATATCATG 852
844 ACCAATGAGTAAAGCGACATTTGCTGATGAGCTATGATGCTGACGAAAT--- 900
853 TATATCTTACTATTTCTACACAAATTCGACCATATATGACAAATATCAACAA 912
901 TGTGTAATGCTATTTCCAAATGCGAATCGAATCGTTTGAACATTTAAATGAAGC 960
913 ACAGCATATGTAATAAATCTCCATATTTTGTAGTAAAGTATGAAATTTATATC 972
961 GCTAAA-AGAACCAACCTTTTGTGATGAGTAAAGTATGAAACAAACAAATC 1019
973 ACTCAATTTCCAG---TGAATAGCAGAAAGAAATTTTACATTAATCCAAATTTT 1028
1020 TTGGAATGCGAAGTATTTGAAGAGAAATATAGTGTGCAAAAGCTTGGTGG 1079
1029 TACTCTGCTGCTCAAGTTTGTCTAAAGAAATTCGAAACATCTTCTTGTATATGTC 1088
1080 TTTCTTGTGACCAAGTATGCGAAAGAGTTCGCTGCTGCTGCAACAACTTTGTC 1139
1089 TTGCTGATTAATAAAGACTGTGATCATTTGATGAAACAAATAGATATGATGTC 1148
1140 CGAGCTGTAAGATTTATCAGCTGCTGAATTTGTTCAAGTCTTTCATCATTAAT 1199
1149 TAAAAATTAATACCATCAAGATTTATAGGTGACAGATTTGAGCTGCAATTAAGACA 1208
1200 TGGCTATTTATAGAAATGATGATGATGTCAGAAATTTGTCGCGCTGAAATAA 1259
1209 TAATGCAATTTGCTGTAATGATGCTGAAATAATTTAGTATTAATGCTGTTGCTA 1268
1260 TTATGCACTAGTGGGAGGAAATTTCTGACGCTTGTGTTAGCGATTAATGCTAAAGCG 1319
1269 CTGTTATTAAGTCAAGTAAATGAAATTAACACTATATATAGCAAAAAATCATTTCA 1328
1320 CACATAAGACTCGGGGATGCAAAATCATCTGCTCGCTGCTGCTGCTGCTGCTGCT 1373
1329 TAGATCTTCAATTAATTTGTCATCATGCTGCTGAGAGTCTTAATTAACATGACGA 1388
1374 CTATTCGCAACATTTTACGAGTGTACAGATTTGCTGATCTTAATCTGTTACTA 1433
1389 CAGAACCTTACGCAATATGCTTTTGGACTAGAAAATAGAAAAGGTAGAAAATTAATA 1448
1434 GTGTGCAATTCAGTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1493
1449 CATTAATAGATCAACCTTAAGCTTGTGAGAGAAACAGTCTGTAATAACAGCATGAT 1508
1494 AAGATTAAGAAATTAAGGCAATGCTGTTGAGGCTTGTGAGACGCTTAAGCAGTACATG 1553
1509 CATTAAGAAAAATTAATGTAAGAACTACAAATTTGCAATCTATTTAACTTTATTAAC 1568
1554 GTTGGGCGAAAAAATGATATGATGCAATTAACCAATGATTTACGGAATTTAT 1613

QY 1569 ATGGAATTA 1579
DB 1614 TCGAAATTA 1624

RESULT 31
US-10-311-455-1280/c
Sequence 1280, Application US/10311455
Publication No. US20030143606A1
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining the Cytosine Methylation
FILE REFERENCE: 5013.1014
CURRENT APPLICATION NUMBER: US/10/311.455
PRIOR FILING DATE: 2002-12-16
PRIOR APPLICATION NUMBER: PCT/EP01/07537
PRIOR FILING DATE: 2001-07-02
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 2424
SEQ ID NO 1280
LENGTH: 6175
TYPE: DNA
ORGANISM: Artificial Sequence
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1280

Query Match
Best Local Similarity 45.8%; Pred. No. 2.7e-12;
Matches 800; Conservative 0; Mismatches 910; Indels 35; Gaps 8;

2.9%; Score 139; DB 14; Length 6175;
Best Local Similarity 45.8%; Pred. No. 2.7e-12;
Matches 800; Conservative 0; Mismatches 910; Indels 35; Gaps 8;

740 ACATATTAACCTTACCGAAAAATTTAAATATTTTACCACATATCTTACGAGC 799
6002 ATAAATTAATAAATAATATATTAATTAATTAATTAATTAATTAATTAATTAAT 5943
800 AACATATGACACGACCAATATATGAGAGATTTATCTGACCAACATATGATATATC 859
5942 TATATTAACCAACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5883
860 TTAATCTATCTTACACCAATTAATGCAATATGTAACCAATTAACCAACAGCAT 919
5882 ACGTAAACAAATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5823
920 ATGCTAAATTAATCTCCAAATTAATTTGTAGTAAAGTATCGAAATTTACATCACTCAA 979
5822 TAAATATTAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5767
980 TTTCCTGATTAATGACAGAAATTTTCAATATATATCAATTTTATATCTCTGCT 1039
5766 ATAAACATATTAATAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5707
1040 CCAAGTTTGTATTAAGAAATTTGACAGACATCTTCTGTAATGATGCTGCTGCTGCT 1099
5706 TACAAATTAATTAATAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5647
1100 AATAAGAACTGCTGATCATGATGAAACAAATTAATTAATTAATTAATTAATTAATTAAT 1159
5646 TATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5587
1160 TACATGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1219
5886 TAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5527
1220 GCATGGAATTAATCGCTGGAATAATTTAGTAATTAATTAATTAATTAATTAATTAATTAAT 1279
5526 AACTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5467


```

/ TITLE OF INVENTION: Prokaryotes
/ FILE REFERENCE: ELITRA.011A
/ CURRENT APPLICATION NUMBER: US/09/815,242
/ CURRENT FILING DATE: 2001-03-21
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ NUMBER OF SEQ ID NOS: 14110
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 6968
/ LENGTH: 1008
/ TYPE: DNA
/ ORGANISM: Haemophilus influenzae
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1)...(1008)
/ US-09-815-242-6968

```

Query Match 2.9%; Score 136.8; DB 9; Length 1008;
 Best Local Similarity 50.4%; Pred. No. 2.9e-12;

Matches 476; Conservative 0; Mismatches 447; Indels 21; Gaps 5;

```

QY 620 ATTAGGTAAAAAATGAAATCAGTATTTTACGTGAGATATTTGGCAGCAGCATATGCA 679
DB 4 ATAAGTTGGCAAAACCCATCAGTGTAGTGGGATCTTATGAACTGGCGCTGGCA 63
QY 680 ATTGACGTGAGCAGCATGTATATCATCTTACTTATGGGACGTGATCATAGAAATAT 739
DB 64 ATTACCTTTTCGGAATGTTCTCCACATCTTATGGGGCACAATCCCGCCATATC 123
QY 740 AOCATATAAACCTTACCGAAAAATTTAAATATTTTACCACATATCATCTTACGAC 799
DB 124 GCACAAATGACAGAGAAAGCAAAATTTATGTTTGGCGATGTCATCTTCCGAG 183
QY 800 AACATATATGACAGCATATATAGAGATATATCTGACACAAATACATATGATATATC 859
DB 184 GATCTTCTTTAGAAAGTATCTTGACACAGCAAT---GGATATTTCTCAAGATATTTTA 240
QY 860 TTAATCTATTCCTACACACAAATTAAGCAGCATATGTAACAAATACACAAACAGCAT 919
DB 241 ATCGGTGGCCAGCAGCATGCTTTCGGTGAATTTCTTATAAAAAT-----CAACCGCAC 294
QY 920 ATGTGTAATAATCTCCCAATATTAATTTAGTAGTAAGTATCGAA---ATTACATCATC 976
DB 295 TTTAAAGGCCATCACCGATTATTTGGGCAACAAAGGTTTGGACGTAACACAGGGCGT 354
QY 977 AAATTTCCAGTGAATAGCAGAAATTTTACATATATATCCAAATTTTATATCTCT 1036
DB 355 TTACACAAACCGTAGTGAAGAACAACTCGGAACGCAATATCCAGTGTATCTTCT 414
QY 1037 GGTCAAGTTTGTGTAAGAAATTTGCAGAACATCTTCTGTATATAGTATCTTGGTGT 1096
DB 415 GGGCCAACTTTTGCAAAAGAAATTTAGCGAAGTGTACATCATAGTATTAAGCTTGCGCC 474
QY 1097 GATAATAAAGAACTTG---GTGAATCATTTGATAGAAACAAATTAAGTATGATTTCTA 1153
DB 475 AATAACGAGCAATTCGACGGGAATTTCAAGTCTGATTCATGTTATAGTAAGTTCCGA 534
QY 1154 AATAATATACCATCAAGATATTTATAGTGTACAGATTTGAGCTGATTAAGAAACATTA 1213
DB 535 GTTATATTAATTTGGATATGACTGGCGTTCAACTTGTGTGAGCAATTAATAATGTGATC 594

```

```

QY 1214 GCATTCATGTGGAATAATTCGTGGAAAAAATTTAGTATATATGCTGTGTACTGTT 1273
DB 595 GCATTCATGTGGAATAATTCGTGGAAAAAATTTAGTATATATGCTGTGTACTGTT 1273
QY 1274 ATACTAAGGCAATGATTAATAATTAACACTATATATAGCAAAAAATCATTCATATAGT 1333
DB 655 ATACTAAGGCAATGATTAATAATTAACACTATATATAGCAAAAAATCATTCATATAGT 1333
QY 1334 CTTCATACATTAATGCTGATTCATGATTCATGATTCATGATTCATGATTCATGATTCAT 1393
DB 709 ACCAATACATTTATGGAATGCTGATTCATGATTCATGATTCATGATTCATGATTCAT 1393
QY 1394 CATTCAGCAATATGCTTTGACTAGAAATAGAAAAAGTAGAAATATTAATATCATTA 1453
DB 769 CATTCAGCAATATGCTTTGACTAGAAATAGAAAAAGTAGAAATATTAATATCATTA 1453
QY 1454 ATGATACCAACTTAAGCTTTGATGAGGACCACTGATTAACCACTGATTAATCATTA 1513
DB 829 ATGAAAAATATGAGGCACTGATGAGGCTTTTATATATACAAAAAGCTTATTTGCTT 888
QY 1514 GCAAAAAAATTAATGATGATTAACCACTGATTAACCACTGATTAATCATTA 1557
DB 889 GCAAAAAAATTAATGATGATTAACCACTGATTAACCACTGATTAATCATTA 1557

RESULT 34
US-10-282-122A-22052
/ Sequence 22052, Application No. US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ PRIOR APPLICATION NUMBER: 60/02-20
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: Patent version 3.1
/ SEQ ID NO: 22052
/ LENGTH: 1008
/ TYPE: DNA
/ ORGANISM: Haemophilus influenzae

```

US-10-282-122A-22052

Query Match 2.9%; Score 136.8; DB 12; Length 1008;
Best Local Similarity 50.4%; Pred. No. 2,9e-12;
Matches 476; Conservative 0; Mismatches 447; Indels 21; Gaps 5;

QY 620 ATTAGTAAAAAATGAAAAATGATATTTAGTGCAGAGATCATTTGGCAGCAATAGCA 679
DB 4 ATTAAGTCCCAACACCAATCACTGCTAGTGGGATCTTAGAACTGCGTGCA 63
QY 680 ATTGACGTGCAGACATGATATATCACTATCTATAGGACGCGATCATAGAAATT 739
DB 64 ATTACTTTTCTGCGCATGCTCTCCACTCCTATGCGGACACATCCGCCATATC 123
QY 740 ACACATATTAACATCTACCGAAAAATTTAAATTTTACCACATATCATCTACACAC 799
DB 124 GCACAAATGCAGACAGAACGACAAATATATGTTTTGGCGAGTGCATCTTCTGAG 183
QY 800 AACATATATGCAGACGAAATATAGACAGATATATGACACAAATACATGATATATC 859
DB 184 GATCTTCATTTGAAAGTATCTTGACACAGCAAT--GGAAATCTTCAAGATATTTTA 240
QY 860 TTAACATATCTTACACACATTAAGCAACATATGATACAAATACAAACAGACAT 919
DB 241 ATCGTGTGTCAGACCATGCTTTGCGTAAATCTTATTAAT-----CAACCGAC 294
QY 920 ATGTGTAATAATATCTCATATTTATTTAGTAAAGTATGAA--ATTACATCTC 976
DB 295 TTAACAAACCATACCGATTAATTTGGGCAACAAAGGTTTGGACGTAACAGGGCGT 354
QY 977 AAATTTCCAGTGAATATAGCAGAGAAATTTTACATATTAATCCATTTTATCTCT 1036
DB 355 TTAACGACAAACGATGTTGTAAGACACATCGGACGATATCCATGCTGATCTTCT 414
QY 1037 GGTCCAGTTTGTCTAAGAAATTTGACAACTCTCTCTGATATGATCTGCTGCT 1096
DB 415 GGCACCACTTTTGCAAAAGAAATTTAGCGAGGTCTACCATCACTATTAACCTGCTGCC 474
QY 1097 GATTAATTAAGAACTG--GTGATCATGATAGAAACAAATAGATGATGTTCTAAA 1153
DB 475 AATTAAGCAATTTGGCAGGGAATTTCACTGCTATCTATCTATGATGAAAGTTTCCGA 534
QY 1154 AATTAATACCATCAAGATATTTAGTGTACAGATGAGCTGCTTAAAGACATATT 1213
DB 535 GTTATATTAATTCGATATGATGCTGCGTTCAACTGCTGAGCAATTAATAAATGTGATC 594
QY 1214 GCAATTTGATGGAATATATGCTGGAATAAATTTAGGTAATATGCTGCTAGCTT 1273
DB 595 GCAATTTGCTGGGATTTTCAACGCTATGGGATTTGGCGAAATGCTGCGACAGCTG 654
QY 1274 AATTAAGCAATGATGAAATTTAAACATATATATAGCAAAAAATCATTCATATGAT 1333
DB 655 ATTACTCGAGTATGCAAAATTAACCGCTTAGTATCTC-----GCTGCTGCAAT 708
QY 1334 CTTCAATACCTATTTGCTCATCATGCTGTGAGATCTAATATTAACATGTAACAGAA 1393
DB 709 ACCAATATCATTTATGGAATGCTGATTTGGCGATTTAGTCTACTGACCGATAT 768
QY 1394 CATTCAGCAATATGCTTTGACTAGAAATAGGAAAGTAGAATATATATATACATTA 1453
DB 769 CAATCGCGCAATGACGTTTGTGATTAATGCTAGTAAGGATTAAGATGCCAATGCT 828
QY 1454 ATAGATCAACACTTAAGCTTTTGAAGACAGTACTGTAACCACTATATCATTA 1513
DB 829 ATGGAATAATGCGGACAGTGTAAAGGTTTAAATATACAAAGAACCTATTTGCTT 888
QY 1514 GCAAAAAAATTAATGTAAGAACTACCAATTTGATATATTTTA 1557
DB 889 GCACAAAGCAAGGTGTGGAATGCCGATTAACAGAAATTTTA 932

RESULT 35
US-10-329-960-1

Sequence 1, Application US/10329960
Publication No. US2003009277A1
GENERAL INFORMATION:
APPLICANT: Fleischmann et al.
TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Fra
FILE REFERENCE: PB186PI
CURRENT APPLICATION NUMBER: US/10/329,960
CURRENT FILING DATE: 2003-01-02
PRIOR APPLICATION NUMBER: US 09/643,990
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: US 08/487,429
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: US 08/426,787
NUMBER OF SEQ ID NOS: 1
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1830121
TYPE: DNA
ORGANISM: Haemophilus influenzae
FEATURE:
NAME/KEY: misc feature
LOCATION: (4747)..(4747)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (9921)..(9921)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (10150)..(10150)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (29298)..(29298)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (36543)..(36543)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (36551)..(36551)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (36636)..(36636)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (40810)..(40810)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (44416)..(44416)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a, t, g or c


```
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (10091)..(10091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (11924)..(11924)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
```

```
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (133160)..(133160)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152530)..(152530)
```

Query Match 2.9% Score 136.8; DB 14; Length 1830121;
Best Local Similarity 50.4%; Pred. No. 5.4e-11;
Matches 476; Conservative 0; Mismatches 447; Indels 21; Gaps 5;

```
QY 620 ATTAGTAAATAAATGAAATGATATTAGTGCAGAGATCTTTGGCAGCAATAGCA 679
DB 635171 ATAACTTGCAAAACCAATCACTGTGTAGTGGGATCTTATGAACTGGCGTGA 635230
QY 680 ATTGACGTGACGACGATGATATCACTTAATTAGGGGACGATGATCAATGAAATTT 739
DB 635231 ATTACTTTTTCGCAATGTTCTTCCACTGATGAGGGGCAAAATCCGCCCATATTC 635290
QY 740 ACACATATTAACACTTACCGAAAAATTTAAATATTTACCCACATATCATCTACAGAC 799
DB 635291 GCACAAATGACAGACGAAACGAAATTAATGCTTTTGGCGGATGTCATCTTCCGAG 635350
QY 800 AACATATATGCAACCAAGAAATATAGAGAGATATTAATCTGACAAACATACATGATATTC 859
```

Db 635351 GATCTTCATTGGAAGTATTTGTCACAGCAAT---GGAATATTTCTCAAGATATTTTA 635407
Qy 860 TTAACATATTCCTACACAAATTTAGCACCATATGTATACAAATACAAACAGAGAT 919
Db 635408 ATCGTGTGTCGAAGCATATCTTTCGTGMAATTTCTTATTAATAAT-----CAACCGGAC 635461
Qy 920 ATGTGTAAAAATATCTCCAAATATTTATTTGTATAGTATCGAA---ATTACATCATC 976
Db 635462 TTTAAAGCCCATACCGATTAATTTGGGCAACAAAGTTGGAGCTAACACAGGGCGT 635521
Qy 977 AAATTTCCAGTGAATATAGCAGAGAAATTTTACATATTAATTCATTTTATCTCTCT 1036
Db 635522 TTACTACAAACCGTATGTTGAAGAACACTCGGAACGACATATTCATGCTGATCTTCT 635581
Qy 1037 GGTCCAGTTTGTCTAAAGAAATTCAGAACATCTTCTTGTATATGTCTTGTGCT 1096
Db 635582 GGCCCAACTTTTGCAAAAGAAATTTAGCGAGGTCTACCATGACTATTAACGCTTGCC 635641
Qy 1097 GATTAATTAAGAACTTG---GTGATCATTTGATAGAAACATATAGATGTTCTAAA 1153
Db 635642 AATTAAGAGCAATTCGACCGGAATTTCACTCTGATTTCAATTTAGTAAAGTTTCCGA 635701
Qy 1154 AATAATATACATCAAGATTTATAGTGTACAGATTGAGCTGCATTAAGACATATTT 1213
Db 635702 GTTTATATTAATTCGAGATATGACTGCGCTCAACTGTGTGAGCAATTAATAATGTATC 635761
Qy 1214 GCAATTCATGTGATATATCGCTGAGAAATTTAGTAAATATAGCTGTTGCTACTGTT 1273
Db 635762 GCAATTTGGTGGCGGATTTTCAGACGCTATGAGATTTGGCGCAATGCTCGACAGCATG 635821
Qy 1274 AATACTAAAGCATGAATGAATTAACACATATATATAGCAAAATTCATTCATATGAT 1333
Db 635822 ATTACTCGAGTATTCAGAAATTTACCCGCTTATGCTATCTC-----GCTCGGTGCAAT 635875
Qy 1334 CTTCAATCATTAATGTGTCATCATGTCTGAGATCTAATATTAATATATATATACACAGAA 1393
Db 635876 ACCAATCATTTATGTGAATGCTCTGATTTGGCGATTTAGTGTCTTACCTGACCGATAT 635935
Qy 1394 CATTCACGCAATATGCTTTTGACTAGAAATAGGAAAGGTAGAAATATTAATATATATTA 1453
Db 635936 CAATGCGCGAATGACGCTTTTGATTAATATGCTAGTAAGGATGATGCCAATGAGCT 635995
Qy 1454 ATAGATCAACACTAAGCTTGTGAAGGAACCATATCTGTAACCACTGATATCATTA 1513
Db 635996 ATGGAAAAATATGCGGCACTCTAGAAAGTTTATTAATCAAAAGAAAGCCTATTTGCTT 636055
Qy 1514 GCAAAAAAATTAATGTAGAACTTACCAATTTGCAATATCTATTTA 1557
Db 636056 GCACAAAGACAGTGTGGAATGCCGATTAACAGAACAATTTA 636099

RESULT 36
US-10-329-670-1
Sequence 1, Application US/10329670
Publication No. US20040018503A1
GENERAL INFORMATION:
APPLICANT: Fleischmann et al.
TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Frag
FILE REFERENCE: P186P1
CURRENT FILING DATE: 2002-12-24
PRIOR FILING DATE: US 09/643,990
PRIOR FILING DATE: 2000-08-23
PRIOR FILING DATE: US 08/487,429
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: US 08/426,787
NUMBER OF SEQ ID NOS: 1
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1830121
TYPE: DNA

ORGANISM: Haemophilus influenzae
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4747)..(4747)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (9921)..(9921)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (10150)..(10150)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (29298)..(29298)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36543)..(36543)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36551)..(36551)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44016)..(44016)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c

```
FEATURE:
NAME/KEY: misc feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
```

```
NAME/KEY: misc feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145342)..(145342)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (152530)..(152530)
```

Query Match 2.9%; Score 136.8; DB 15; Length 1830121;
Best Local Similarity 50.4%; Pred. No. 5.4e-11;
Matches 476; Conservative 0; Mismatches 447; Indels 21; Gaps 5;

```
QY 620 ATTAGTAAATAAATGAAATGATATTTAGTGCAGATCTTTGGCAGCAATAGCA 679
DB 635171 ATTAAGTGCAGCAATGATCTGCTGAGGAGTCTTAAGCACTGCGTGGCA 635230
QY 680 ATTCAGTGCAGCAATGATATTCAGTTAATTTAGGAGCGATCATAGAAATAT 739
DB 635231 ATTAAGTGCAGCAATGATCTGCTGAGGAGTCTTAAGCACTGCGTGGCA 635290
QY 740 AACATATGCAACGCAATATAGCAAGATATTCGACACATATCATCTTACAGAC 799
DB 635291 GACCAATGCAACGCAATATAGCAAGATATTCGATCTTACAGACATCTT 635350
QY 800 AACATATGCAACGCAATATAGCAAGATATTCGACACATATCATCTTACAGAC 859
DB 635351 GATCTTCATTTAGAAAGTATTTGACACAGCAAT--GGAATATTCACATATTT 635407
QY 860 TTAATCTTCTTCAACACAAATTAAGCAATATGATGACAAATACACAAAGCAT 919
DB 635408 ATCTGCTGCAACGCAATGCTTTCGTAATCTTAATAAAT-----CAACGCGAC 635461
QY 920 ATGTGTAATAATACCAATATTAATTTGTAAGGATAGAA--ATTGATCATCTC 976
DB 635462 TTAAGGCAATCAGCATTAATTTGGGCAACAAAGTTTGAACGTAACAGGCGT 635521
QY 977 AATTTCCAGTGAATATGACAGAAATTTTACATATTAATTCATTTTACTCTCT 1036
DB 635522 TTAAGGCAATCAGCATTAATTTGGGCAACAAAGTTTGAACGTAACAGGCGT 635581
QY 1037 GATCCAGTTTGTCAAGAAATTTGCAAGAACTCTTGTAGATATGATCTCTGCT 1096
DB 635582 GATCCAGTTTGTCAAGAAATTTGCAAGAACTCTTGTAGATATGATCTCTGCT 635641
QY 1097 GATATTAAGAACTGT--GTGAATCATGATAGAAACATTAAGTATGATGTTCTA 1153
```

Db 635642 AATAAGAGCAATTCGACGGGAATTGAGTCTCGATTCATTGAGTAAGAGTTTCCGA 635701
QY 1154 ATATATACCATCAAGATATTTAGTGTACAGATTGAGCTGCATTAAGACATTAAT 1213
Db 635702 GTTTATATAAATTCGATATGAGCTGCGCTTCACTGTGAGCAATTAAGATGATC 635761
QY 1214 GCAATTCAGTGTGAATATCCCTGGAAAAAATTTAGGTAATATAGTGTGTACTGTT 1273
Db 635762 GCAATTCGAGGGATTTTCACACGGTATGGGATTTGGCGAAATGCTCGACAGCACTG 635821
QY 1274 ATATCTAAGGCAATGATTAATTAACACTATATAGCAAAAAATCATTCATATGAT 1333
Db 635822 ATTTACCGGATATGACAAATTAACCGCTTAGATATCTC-----GCTGCGCAAT 635875
QY 1334 CTTCATACATTAATTTGCTCATCATGCTTTGAGAGTCTAATATATACATGTACAAAGAA 1393
Db 635876 ACCAATACATTTATGGGAATGCTGATTTGGCGATTTAGTCTACTTGCACCGATAT 635935
QY 1394 CATTCACGCAATATGGCTTTTGGACTAGAAATAGAAAAAGTAGAATATATATACATTA 1453
Db 635936 CAATCGCGCAATCGACGTTTGTGATTAATGCTAGTAAAGATTTAGATGCCCAATGGCT 635995
QY 1454 ATAGATCAACACTAAAGCTTTGAGAGAACAGTACTGTAAACCATGATATCATA 1513
Db 635996 ATGGAAAAATATGGGCAATGCTAGAAAGTTTATTAATACAAAGAAAGCCATTTGCTT 636055
QY 1514 GCAAAAAAATTAATGTAGAACTACCAATTTGCATATCTATTTA 1557
Db 636056 GCACAAAGACAAGGTGTGAAATGCCGATTAACAGAAATTTA 636099

RESULT 37

US-10-311-455-2213/c
Sequence 2213, Application US/10311455
Publicatation No. US20030143606A1
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENBROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining the Cytosine Methylation
FILE REFERENCE: 5013.1014
CURRENT APPLICATION NUMBER: US/10/311,455
CURRENT FILING DATE: 2002-12-16
PRIOR APPLICATION NUMBER: PCT/EP01/07537
PRIOR FILING DATE: 2001-07-02
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 2424
SEQ ID NO 2213
LENGTH: 11691
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-2213

Query Match 2.9%; Score 136.2; DB 14; Length 11691;
Best Local Similarity 43.1%; Pred No. 9.5e-12;
Matches 1193; Conservative 0; Mismatches 1543; Indels 31; Gaps 10;

QY 256 ATAAATCTATCAAGTCTCTTACATAGAAACACTGTATTCAGAAATACCTCTCTCT 315
Db 11118 ATAAATCTATCTATCTTATTAATAAATAAACAATCTTTAATCTCAGCTATATAT 11059
QY 316 TGTTCATCATATTTTCCATATTTTACATATATATATAGTCTCTTATTTGCAATACAT 375
Db 11058 TTTTCTTAAATATTTAACTCTTATTTTCAATACAAATTTTCTATCTTAAATTAAC 10999
QY 376 ATTAATATCAATCAAAACAAATTAATAGTACTATCTAATTAATTAATTAATTAAT 435

Db 10998 ATAAATTAATAAACTCACTAAATTTTAAACATTTTATTTTTCATTAATCCAAAT 10939
QY 436 TCTTGAAGAAAGCTCAATATATATATAGACACTTAATATATATATTTCAATTTA 495
Db 10938 AAATTAATTAATAAATCTAATATTAATAAATATCTATTTATTTTCAATTAATAAAT 10879
QY 496 TTATACACTCATATTTTATAGATATATATATATATATATATATATATATATATAT 555
Db 10878 AAAATTTATTAATTAATTAACCAAAATTAATTAATTAATTAATTAATTAATTA 10819
QY 556 ATTTCTTAACATATATATATATATATATATATATATATATATATATATATATAT 615
Db 10818 AATATTTCTACATATTAATCTATATATATATATATATATATATATATATATAT 10759
QY 616 ATAAATTAAGTAAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 675
Db 10758 CTCTTATA-AAAAATTTAAACCTTAATTAATTAATTAATTAATTAATTAATTA 10700
QY 676 AGCAATTCAGCTGACACATGATATATCACTTAATTAATTAATTAATTAATTAATTA 735
Db 10699 AACATATATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 10640
QY 736 TATTAACATATTAACACTTACCGAAATTAATTAATTAATTAATTAATTAATTAATTA 795
Db 10639 ACATTTCTAAATATATTTACTTTAAACACACACAAAAAATTAATTAATTAATTTTC 10580
QY 796 AGCAACATATATATGACACACACATATAGACAGATATATGACACACATATATAT 855
Db 10579 ACATTAATATATTTTAAACATTAATTAATTAATTAATTAATTAATTAATTAATTA 10521
QY 856 TATCTTATATCTTCTTACACACAAATTAATTAATTAATTAATTAATTAATTAATTA 915
Db 10520 AAAAATTAATTAATTTAAATTTCTATTAATTAATTAATTAATTAATTTTAAACAA 10461
QY 916 GCATATGATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 975
Db 10460 ATATATATATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 10401
QY 976 CAAATTTCCAGGAAATGACAGAAATTTTACATATATTAATTAATTTTATCTGTC 1035
Db 10400 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 10341
QY 1036 TGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 10283
Db 10340 AAAATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 10283
QY 1096 TGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1155
Db 10282 AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 10223
QY 1156 AATATACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1215
Db 10222 TTAACATTTTCACTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 10163
QY 1216 AATGATGATGATATATGCTGAAAAATTTAGGTAATTAATGCTGTGCTACTGAT 1275
Db 10162 ATCTCACTACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 10103
QY 1276 AATTAAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1335
Db 10102 ATCTCACTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 10043
QY 1336 TCAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1390
Db 10042 TAAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 9983
QY 1391 GAACATTCAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1450
Db 9982 TAAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 9923
QY 1451 TTAATGATCAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1510
Db 9922 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9863

| | | | | | | | |
|----|------|-----------------------------------------------------------|------------------|-----------------------|---------|----|------|
| QY | 1511 | TTAGCAAAAAA | CTTAATCTAGACATCA | CCAAATTCGATCTATCTATTA | CTTATTA | CT | 1570 |
| Db | 9862 | ATACATATACATTTATATATATCTTCAAAATTTTCTMAAATATATTCATTTACAAAT | | | | | 9802 |
| QY | 1571 | -GAGAAATATCATAGATTAAGCAATATCAACATATTAATCTTAGTCATATCTTTG | | | | | 1623 |
| Db | 9802 | ATAAATATTAATTTTACCAACATTAATTTTMAAATTTTMAACCAATACATATAC | | | | | 9743 |
| QY | 1630 | ATAATTCACAACAATGATTAATTAACGGTCAAAATATAGATTAACAAGTTCAACCAT | | | | | 1689 |
| Db | 9742 | TAAATCCAAATTAATCTATATATATATACAAATTAATTTATTCAAAATTAATTA | | | | | 9682 |
| QY | 1690 | ATATMAAAGCGTCAATACCATGCTTAACATCCAACTAAGATGTAACATATTACAT | | | | | 1749 |
| Db | 9682 | TATTAATMAACATTAACAATTTMAATATTCATTTCTTTTTCMAATCCTTATCTA | | | | | 9623 |
| QY | 1750 | AACAACATTAAGTATATACCTTAATTTCCATCTATATATATTAACAATTCCTATCC | | | | | 1809 |
| Db | 9622 | ACATTTCAATTAATTAATTAACATATCTAAATAGATTAATTTACTTMAAATCTCCAA | | | | | 9566 |
| QY | 1810 | TATTCATTAATCAATGTTATTCAGATTAATTAATGCAATTAACCTTTACTAT | | | | | 1865 |
| Db | 9562 | CATATMAAATTTCTAATCCTATTAATATMAACATTCATTAATTAATTTTCCATTT | | | | | 9503 |
| QY | 1870 | TATAAATAATGTTGTCACCTTTTCACATCATCATGATATATATGAACATATTT | | | | | 1929 |
| Db | 9502 | TCTTAATTAATTAACAAATCTTTTATATMAAATTCATACAAATATATATTAAT | | | | | 9443 |
| QY | 1930 | ACATMAATTAATTAATTAAGAACCAATACATAT- - - - -TAAATTAATAAAC | | | | | 1982 |
| Db | 9442 | TACTMAAATATCTTTCTMAAATCGMAAATATATMAACMAAATATCTATTAACAC | | | | | 9383 |
| QY | 1983 | CTATTTACGTATATTTACATAGTC- - -CTGACATATMAATCTATCCTCTG | | | | | 2039 |
| Db | 9382 | ATTTCTCATACATAATATATATATTCATTTTMAAATTCATTAATTTTCT | | | | | 9323 |
| QY | 2040 | ACATATTTCAACATMAATATGATTAACGTTTTTTATGCCATAGTCTTATMAA | | | | | 2099 |
| Db | 9322 | AAATTAACATTTAATTAATTAATTAATACATATTAATCTTTTATCTTTACTTAA | | | | | 9263 |
| QY | 2100 | ATTGCTAATTAACCAAAAGTTAAATATACATCTCGCAACATATGAAGAATCA | | | | | 2159 |
| Db | 9262 | ATTTAAATATAC- - -AACTAATTTAAATTTTATTAATAAACAATAATTAAT | | | | | 9207 |
| QY | 2160 | ATACAGCTCAATTAATTAACAAATTTCTATATATACACCTACTTAACATAGAC | | | | | 2219 |
| Db | 9206 | ATMAAACCTTAATTAATTAATCTTAACCTAACACCAACCAATATATCATCAAT | | | | | 9147 |
| QY | 2220 | CAATGATATCATCTCTAAGATATCCCTACAGATGATATCAATATTAAGAAA | | | | | 2279 |
| Db | 9146 | ACGTATGTAATTTCTAACTCTAATCATATTTATTTTAAACCAATMAAATATACC | | | | | 9087 |
| QY | 2280 | ATCGTATGATTTATTAACAATTAACAATTTCAACTTACACACAAACCTAAAGT | | | | | 2339 |
| Db | 9086 | TATCACTCTTAATTAATTAACCCCAAAATTCATACAAACCAAAATTAAT | | | | | 9027 |
| QY | 2340 | CTATTAATACAAATCAAGATTAATATATCTTACCAACAACATAGTATCAATTCAT | | | | | 2399 |
| Db | 9026 | CTAACATTTTATTTTCAACTATATATACATATTCATATTAATTTTTCAT | | | | | 8967 |
| QY | 2400 | AATACATATATCTTTTAAAAAGATGACTTCACAATTCATACATATATTAATCAT | | | | | 2459 |
| Db | 8966 | ATAATATATATTAATCACTCCCAATTTACATTAACCTTMAAATTAACAAATAT | | | | | 8907 |
| QY | 2460 | GCAATGTTTATTCAGAAACATTAACAATCCAAATTTTCAATTTACGGGTATCGT | | | | | 2519 |
| Db | 8906 | TAAATATATTTCAAAATTAACCTTAATTAATATCTATATCTACAAACACAAAA | | | | | 8847 |
| QY | 2530 | AACGTTCAATCCAAACAAATTCAGAGTGTGATCATTAATCTCCAAACGAGATTA | | | | | 2579 |
| Db | 8846 | CTTAACCTTAATCACTACATATATATTCAAAAATTCATTTTTCATTAATCATATA | | | | | 8787 |

```

QY 2580 GAGTTTAAATCTGCTCTTCAATTCCTCAGAGCTTACGTAATGTAAGAATAA 2639
Db 8786 ATTTTATCAATATCTAATTAATAATTTTCTACTTTTATCTTAATATTTAATA 8727
QY 2640 CTTTACGACTTAAACATCTAATTAAGCGAATTAGTACTATCTACTGTTAAATCTT 2699
Db 8726 TTTTATTTCAAACTAAATCTTTTTCATTTTCTACCTTTAAATATA-----ATCACC 8672
QY 2700 TTGCTGATCATCTCCCTTAGGAGATTTCCAAATTAATTCATCTATAACATTAAGTA 2759
Db 8671 CTAATTCATCTACTATTAATAATTTTACCAAAATATAATACATACCAAAAAC 8612
QY 2760 ATGCTCTTAAGTACCTGGCCCATCATATGACGTGCTATTA--AAGCTTCATCACT 2817
Db 8611 AAAACCTAATTAACCTCCCTAATAATAATTAACCTTAACAAATTAATTAATCTCACT 8552
QY 2818 TAATACATGAGTACACAAACCTATTTTTTTTACAGATGCTATTCATAAATACCTACACAA 2877
Db 8551 AAATATATACATACCTTACATTTTATATATATACCTTTATATATATATACACACCT 8492
QY 2878 TTCGGATATCTGTACAAATTTCTACTAAAGAAGTACTCTAATTGAAAGACATTTAA 2937
Db 8491 ACTAAATTAACATTTATATCTTCACTTTTATTAATTAATTAATTAATTAATTA 8432
QY 2938 TCTATATATTAATCCCTACAAACCTACAGCTTTTACTCTCACTTCAATCTTTGGA 2997
Db 8431 TAAATTAATCTTAATTAATTCATACCAATTAATTAATTAATTAATTAATTAATTTTA 8372
QY 2998 AGAAGAC 3004
Db 8371 CAACAC 8365

RESULT 38
US-10-311-455-1999/c
; Sequence 1999, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIPENBROCK, Christian
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining the Methylation of Cytosine Residues in DNA
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311.455
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1999
; LENGTH: 17594
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1999

Query Match 2.98; Score 136; DB 14; Length 17594;
Best Local Similarity 45.0%; Pred. No. 1,2e-11;
Matches 996; Conservative 0; Mismatches 1190; Indels 26; Gaps 12;

QY 345 AATATCATAGAGTCTTTATTCGACATACATATTAAATCAATCAAAACAAATATAGT 404
Db 7870 AATATTAATAAACTTAATAAAATCAAAATATATATATATATATATATATATATAT 7811
QY 405 ACTATCTAATTAATTAATAAAATCCTCAATAGTCTTTGACAGAACTACAAATTAATGAG 464
Db 7810 ATTAATAATAAAATTAATAATTAATTAATAATTAATTAATTAATAATACTTACAAAT 7751

```

| | | | |
|----|------|---------------------------------------------------------------|------|
| QY | 465 | ACACCTACTCTTTTAAATATATTTGCAATTTATTAATACAGCATGATTTTGGCATTA | 524 |
| Db | 7750 | TAAATAACGCAAAAATATATAATTAATCTTAAATTTCTACATATACTAAATTA | 7691 |
| QY | 525 | TCACACTGCACTGATTAACAATAGTAACAATTCTTTACACATGATATATACATA | 584 |
| Db | 7680 | ACAAATAAATTCATTAATTAATAAACCTAAAAATCTATATT---TAAAAAATACATAAA | 7638 |
| QY | 585 | ATATATATTAATTTTGACTGGTTTCTGATATTAATTAAGTAATAAATGAAATCAGTA | 644 |
| Db | 7634 | TAAAAAATAATCAACCAAAATATTCGATTTAAAAAATATTTTAATTTAAACAATCTTTA | 7575 |
| QY | 645 | TTTTAGTGCGAGATCATTTGGCAACGCAATAGCAATTGCACTGCAGCAATGATAT | 704 |
| Db | 7574 | TTATATCTAAAAATCTATATCTAAATPAACAATTAACAAAAAATTTAAAAAACCAATPA | 7515 |
| QY | 705 | CAGTTAACTTATGGGAGCGATCTATGAAATATATACATATTAACCTTAACGATAA | 764 |
| Db | 7514 | AAATATCTATATTAATAAATATATAAATTAATAAATAATCACTAACATATCTTTAAA | 7455 |
| QY | 765 | ATTAAATATTTTACCACATATCAATCAACGACAACATATATGCAACCGAATATAG | 824 |
| Db | 7454 | CCACTAAATTAATAATTAATCAACTAAAAAATAAATAAATAAATAAATAAATAAATAA | 7389 |
| QY | 835 | ACGAAGTATTAATCTGACACACATACATGATATATCTTACTATTCCTCACACACATTA | 884 |
| Db | 7384 | TCTCAAAACAAAACCTTAATAAATTTCAATATCTTAATAATTAATATCAAAAAACCAA | 7335 |
| QY | 885 | GCACATATGTAACAATAATACACACAAACGACATATGTGTAATAATATCTCAATTTTA | 944 |
| Db | 7334 | ACAAAACCTAAAAAACAACCTAATTAATTAATAAATAAATAAATAAATAAATAAATAA | 7275 |
| QY | 945 | TTTGTAGTAAAGTATCGAAATTTACATCACTCAATTTCCAGTGAATATACAGAGAAA | 1004 |
| Db | 7274 | AACCAAAAAAATAACAATCTAAAAAATTCGTTAATTAATTAATCTTAATTAATAA | 7215 |
| QY | 1005 | TTTTCAATATTAATCCAAATTTTATACTCTCGTGCCAAAGTTTGTCTAAAGAAATGAG | 1064 |
| Db | 7214 | ATAAATTAATAAATAAATAATACATTAATCTCTC---AAATCAACAATATTAATAATTTT | 7155 |
| QY | 1065 | AACATCTTCTGTAGTATAGTACTGTGCGTGTGATATTAAGAACTTGTGAATCATTA | 1124 |
| Db | 7158 | AAAACTTTAACATATACATATAAATAATATATATATTAATAAACAACCTAATAATTA | 7099 |
| QY | 1125 | TAGAACAATAGTATGATGATGTTCTAAATAATATTAATCAATCAAAATTTATAGGTAC | 1184 |
| Db | 7098 | AAATTAATTAATAATTAATAAATAAATCAATTAACAATATAGAACAA---TATCTAAAAAT | 7041 |
| QY | 1185 | AGATTGAGCTGCAATTAAGAACATTAATGCAATGTCATGTGGAATATCGCTGGAAAA | 1244 |
| Db | 7040 | TTACTATTAATAATTAATAATTAATAAATAAATAAATAAATAAATAAATCTTAATTA | 6981 |
| QY | 1245 | ATTGATTAATTAATCGTGTGCTACTGTATTAATTAATTAAGCAATGAATTTAAACAC | 1304 |
| Db | 6980 | AAATCTAAATTAAT---ATTATTAATTAACAATAAATTAATTAATTAACAATAAATAA | 6922 |
| QY | 1305 | TATATATAGCAAAAAATCAATCAATAGATCTTCATACATTAATGTGTCATCATGCTTG | 1364 |
| Db | 6921 | AAAAATATTAATTTAAATTAATTAATAAATAAATAAATAAATAAATAAATAAATAA | 6862 |
| QY | 1365 | GAGATCTAATTAATTAACATGTAACAAGAACTTCAACGAATATGAGCT---TTTGAAGTA | 1423 |
| Db | 6861 | TAACTTCACTCATCAATTTTAATCTTTCTTAATAATTAACAATAATTTCTACATTA | 6802 |
| QY | 1424 | ATAGAAAAGGTAGAAATATAAATACATTAATAGATCAAACTTAAGCTTGTGAAGGA | 1483 |
| Db | 6801 | CATATTAATCAATTTATTTTAAATTAATTTTCAATTTCAAAAAATACTAAATCTA | 6742 |
| QY | 1484 | ACCACTACTGTAACCAACGATATCATTAAGCAAAAAATAATATGA---GAACTACCAAT | 1542 |
| Db | 6741 | ACAATCTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA | 6682 |
| QY | 1543 | TTGCATATCTATTTGCAATTTTATTAAGATGAAGATATATCACTAGTAAAGCAATATCAA | 1602 |

| | | | |
|----|------|----------------------------------------------------------------|-------|
| Db | 6661 | TAAGCTAAACGTTAAATTTCTACCAATAAAGCTATACAAATTAAATAATATATTAATT | 66222 |
| QY | 1603 | CATATTTATCTTAGTCTATC-ATCCTTGATTAATTCAACACATGATPAATPAACGGTCA | 1661 |
| Db | 6621 | TATATATAAATTAATTAACATAAATATGTAACCTTAATCTTTCAAAAATTAATTAATAA | 6562 |
| QY | 1662 | AAAATATGATTAACAAGTTGACACACATATTAATAAGACGTCAAATACCATGGTTAAC | 1721 |
| Db | 6561 | AAATTAACAAAAAATAATACACAAAAAACAATAAACCATAACATCTCAAA | 6502 |
| QY | 1722 | ATCCAACTAAGTATGTAAACATATTTACATAACAACTTAAGTATTAACCTTAATTTCTGAG | 1781 |
| Db | 6501 | ATACAGCTATATATAAAGCTAAACCAAAAAAACCCTTAATAATATATTAATCTATACAA | 6442 |
| QY | 1782 | CTAATTAATPAACAAATTCCTATCTCTATCTCAATTAACCAATTTGTTATACATTAATA | 1841 |
| Db | 6441 | ATATTCATTTCCCTTACCATAT--TCATCTATACTAAATTAATTAATTAATCTT | 6384 |
| QY | 1842 | AATAAGTCAAAATTAACCTTTACTTATTAATAAATATATATGTTGACACTTTTCACT | 1901 |
| Db | 6383 | AATATTTCAAAATTTAACTTACATA--AAAAATCTTATTTACACCACTTCAATAT | 6326 |
| QY | 1902 | ACATCATGTATATATGAACATAATTTTAACTAAATTAACCTAATTAATAAGAACATATCA | 1961 |
| Db | 6325 | ATATTTAATATTTCCCTTAACCTTACCTATTAACATATTAACCTTAATTTCAATTACAAAAA | 6266 |
| QY | 1962 | TATATTAATTAAGTAAACACCTTATTAAGTACAGTAATTAACATAGATGTC--TGACA | 2018 |
| Db | 6265 | TACATTTATTAATAAATACTACATATATTTCTTAATTTATTAACATTTACCATATACA | 6206 |
| QY | 2019 | TAAATTAATTCATCATCCTTGACATATTTCAACATTAATAATGATTAACCTGTTTTTTA | 2078 |
| Db | 6205 | TAACATATATATAATATCTACGTAAATACAAAAATCAAAAAAATAATTCACAAA | 6146 |
| QY | 2079 | TGCCCATTTAGTCTTAATTAATAATGCTAATTAACCAAAAAAGTTAAATACACTTCCTG | 2138 |
| Db | 6145 | CCACAATTAACCTTTATTAATAA---CTTAATTTCCATTAACCAAAAAATPATTAATAA | 6090 |
| QY | 2139 | CAACATATGAGAGATACATACACGTCAAAATTAATTAACCAAAATTTCTAATATAC | 2198 |
| Db | 6089 | CTAAATATATTTCTAAATTAATAATTAATAACCAATACTATTAATATTTTTTAAACCTA | 6030 |
| QY | 2199 | ACCTACTCTAAACATAGGAACCAATAGTATATCTAC-CTAAGTAAGATATCCTACAC | 2257 |
| Db | 6029 | AGTATATTCAAACATTTACTTTATCCTATTTATATCTCAAAAAACCTATTAATAACAA | 5970 |
| QY | 2258 | ATGATTAATAATTAATAAGAAAAATCGTATGATATTTATATACATTAACAAATTTCCAC | 2317 |
| Db | 5969 | TACTATCAACATCCTTATTTTAAATTAATAAATTAATAATTAATTAATTAATTAATTAAT | 5910 |
| QY | 2318 | TTCAACAACAAACCTCAAAGTTCTATTAATAACAAATTCAGAAATTAATATATATCTTCA | 2377 |
| Db | 5909 | AACAAACCAAAATTTTAAACTTTAAATCTTAATTAATTAATAAATCTTAACATTAATAA | 5850 |
| QY | 2378 | ACAATATGTTATGCAATTCATTAATATACATATATCCTTTTAAATAAGTACTCAAT | 2437 |
| Db | 5849 | AAAAACATTCACAAATTAACCTCCACACTTTAAACACACATTTAAATTAACCTTAATAT | 5790 |
| QY | 2438 | TCATATCATTAATTAATTAATGATGCAATGTTATTCAGAAAATTTACAAATTCACAAATTC | 2497 |
| Db | 5789 | TCCCTCCAAATTAATTAATAATTAATATCATCTTTTAACTTAATAATTAATAACAAAT | 5730 |
| QY | 2498 | TTCAATTTACGGGTATATGCTGAACGTTCCATTCACAAATTCAGACAGTTTC | 2545 |
| Db | 5729 | ACTATATAAAAACTTAACACTACTAAATATCTAATATAAAAAATCTTAACACTAC | 5678 |

RESULT 39
US-10-311-455-1460/c
; Sequence 1460, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:

APPLICANT: OLEX, Alexander
APPLICANT: PIRENBERG, Christian
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detect
FILE REFERENCE: 5013.1014
CURRENT APPLICATION NUMBER: US/10/311.455
CURRENT FILING DATE: 2002-12-16
PRIOR APPLICATION NUMBER: PCT/EP01/07537
PRIOR FILING DATE: 2001-07-02
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 2424
SEQ ID NO 1460
LENGTH: 6831
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1460
Query Match
Best Local Similarity 2.9%; Score 135.8; DB 14; Length 6831;
Matches 1808; Conservative 0; Mismatches 2322; Indels 89; Gaps 19;
QY 239 AAGAACTCCCTCTCTGTTCCATCCATATTTTCATTTTACATATCATAGAGTT 358
DB 6580 AATTAACCTTATATATATATTTCTATCTTAAATATACATCAAAAACAAATTTT 6521
QY 359 CTTATTCGACATCATATTTAAATCATCAAAACAAATATATGT-----ACATTC 410
DB 6520 ATCTTTACACAAATATATCTTAATCTTAATATTTACATATATTTTACTATTTCTAATA 6461
QY 411 TAATTAATTAATAAATCCCTCATAGTCTTTAGCAAGAACTCAATATATATAGACACTT 470
DB 6460 CCAAAATTAACATATATTAACAAACCTTACCTCATCAATATATATATTAATTAATAA 6401
QY 471 AACTTATTAATATATTTCAAAATTTATTTACTACTCATATTTAGCATATATTCAT 530
DB 6400 AAATTAACAAATTAATAAATAAACAACATCTAATTAACAAACAAATTTTCAATATAT 6341
QY 531 CTGACATGATTAACATATGTAACATTTCTTAACATGATATATATATATATATA 590
DB 6340 AATATATATATATATATTTCTTAACAAATCTTAATAAACAATTAATAAATAAATAA 6281
QY 591 TATATATTTTGAATGTTCTTGAATATATATATATATATATATATATATATATAT 650
DB 6280 AATTAATTTTAA-----TACTAATTTTAAATAAATACTTCTTAATAAATAAATAA 6225
QY 651 GTGCGATCATTTGGCAGACATATAGATTTGACATGTCAGCAGCATATATATCTGTTA 710
DB 6224 AAAAAAACAATATGTAATCTTAACGAAATATCTTAAATAAACAACAACAATATACA 6165
QY 711 ACTATGAGGAGTATGATATGAAATATATACATATATTAACATCTTACGAAAAATTTAA 770
DB 6164 TCTCTTTTAAATTAATAAATACTTAACAAATCTAATAAATAAACA--ACATATATATATA 6108
QY 771 AATATTTTACCAATATCATATACACACAAATATATGCAACGACAAATATATAGAGAG 830
DB 6107 AAAAAATATATATATATATATATATATATATATATATATATATATATATATATAT 6048
QY 831 TATATCTGACACATATATATATATATATATATATATATATATATATATATATATAT 890
DB 6047 TATATATATATATATATATATATATATATATATATATATATATATATATATATAT 5988
QY 891 TATGTAACAAATATACACAAACAACATATATATATATATATATATATATATATATAT 950
DB 5987 ATTTTAAACAAATATATATATATATATATATATATATATATATATATATATATATAT 5928
QY 951 GTAAGGATGAAATATATATATATATATATATATATATATATATATATATATATATAT 1002

DB 5927 ACTAAAAAACAATAATATATATATATATATATATATATATATATATATATATATAT 5868
QY 1003 AATTTTACATATATATATATATATATATATATATATATATATATATATATATATAT 1062
DB 5867 ATTTAAAAAATAATATATATATATATATATATATATATATATATATATATATATAT 5808
QY 1063 AGAATCTCTCTCTGATATATATATATATATATATATATATATATATATATATATAT 1122
DB 5807 AGCTTTTAAATATATATATATATATATATATATATATATATATATATATATATAT 5748
QY 1123 GATGAAACATATATATATATATATATATATATATATATATATATATATATATAT 1182
DB 5747 AATCTATATATATATATATATATATATATATATATATATATATATATATATAT 5688
QY 1183 ACGATTTGAGCTGATATATATATATATATATATATATATATATATATATATATAT 1242
DB 5687 ATATATATATATATATATATATATATATATATATATATATATATATATATATAT 5628
QY 1243 AATTTTGTATATATATATATATATATATATATATATATATATATATATATATAT 1301
DB 5627 ATTAATCTATATATATATATATATATATATATATATATATATATATATATATAT 5568
QY 1302 CACTATATATATATATATATATATATATATATATATATATATATATATATATAT 1360
DB 5567 AATATATATATATATATATATATATATATATATATATATATATATATATATATAT 5508
QY 1361 CTGAGATCTATATATATATATATATATATATATATATATATATATATATATATAT 1420
DB 5507 ATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 5448
QY 1421 GAATATGAAATATATATATATATATATATATATATATATATATATATATATAT 1477
DB 5447 TATCTGACAAATATATATATATATATATATATATATATATATATATATATATAT 5388
QY 1478 GAAGAACAGATATATATATATATATATATATATATATATATATATATATATAT 1537
DB 5387 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 5328
QY 1538 CCAATTTGATATATATATATATATATATATATATATATATATATATATATATAT 1597
DB 5327 CTATATATATATATATATATATATATATATATATATATATATATATATATATAT 5268
QY 1598 TCAATATATATATATATATATATATATATATATATATATATATATATATATAT 1637
DB 5267 TCTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 5208
QY 1638 AACACATGATATATATATATATATATATATATATATATATATATATATATATAT 1697
DB 5207 ACTACACCTTAACTCTCTAAATATATATATATATATATATATATATATATATATAT 5148
QY 1698 AGCAGTCAATATATATATATATATATATATATATATATATATATATATATATAT 1757
DB 5147 AAAAACTAAAGCGTATCTACACACCAATATATATATATATATATATATATATATAT 5088
QY 1758 TAAGTATATATATATATATATATATATATATATATATATATATATATATATATAT 1817
DB 5087 TTCACATATATATATATATATATATATATATATATATATATATATATATATATAT 5028
QY 1818 TTAACCAATTTTATATATATATATATATATATATATATATATATATATATATAT 1877
DB 5027 TTCCAAATATATATATATATATATATATATATATATATATATATATATATATATAT 4969
QY 1878 AATATGTTGACACATTTTATATATATATATATATATATATATATATATATATATATAT 1937
DB 4968 AATATATATATATATATATATATATATATATATATATATATATATATATATATAT 4909
QY 1938 TAACTATATATATATATATATATATATATATATATATATATATATATATATATAT 1997
DB 4908 CTCAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 4849
QY 1998 TATTAATATATATATATATATATATATATATATATATATATATATATATATATAT 2057
DB 4848 AAACATTAATATATATATATATATATATATATATATATATATATATATATATATAT 4789

QY 2058 ATATGATAAAGTGTGTTTATGCCCCATAGTCTTAATATTAATGCTATTATACCAAA 2117
Db 4788 ACATTTCCTACATCTCTAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTCAT 4729
QY 2118 AGGTAAATATACATCTCTGCAACATATGAGAAAGATACATACGCTCAATTAAT 2177
Db 4728 ATTTCTAATAAA---ACTATTAATAAATAACATATATCTTCATATATTAATCAAT 4673
QY 2178 AAAAATTTCTAATATTAACCACTACTTAACATAGGAAACCAATAGATATCAATCT 2237
Db 4672 AAAAAAATAAATAAATAAACAATAATTAAT---ACAAATTAATTAATCTTAATTAATCT 4616
QY 2238 AAGTAAGATATCCCTACAGATGTATCAAAATTAATAAGAAATAATGTTATGATATTAAT 2297
Db 4615 AAAAAACCACTAATCTATTTTATACAAATTAATAAATAACATCAATCAATCACTCAAC 4556
QY 2298 ACAATTAACAATTTCCAACTTCAACAAACCAAACTCAAGTTCTATTAACAAATTCGA 2357
Db 4555 AAATTAATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4496
QY 2358 GAATTAATTAATCTTCAACAAACATATGATTAATTAATTAATTAATTAATTAATTAAT 2417
Db 4495 TACTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4436
QY 2418 TAAATAGATGCTCAATTCATACATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2477
Db 4435 AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4379
QY 2478 CATTAACAATTCAGAAATTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2537
Db 4378 ATTCATCTTAACTTCTTCAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4319
QY 2538 ATTCAAGATGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2597
Db 4318 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4259
QY 2598 TTTCAAATTTCTGAGAGCTTTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2657
Db 4258 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4199
QY 2658 CATTAATAGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2717
Db 4198 ATTACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4139
QY 2718 TAGAGATTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2775
Db 4138 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4079
QY 2776 TGGCCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2835
Db 4078 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4019
QY 2836 ACCATTAATTTTCAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2895
Db 4018 ACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3959
QY 2896 ATATTTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2955
Db 3958 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3899
QY 2956 ACAAAATCTACAGCTTTTCACTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3015
Db 3898 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3835
QY 3016 CAATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3070
Db 3838 CCGGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3779
QY 3071 AATTAATCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3130
Db 3778 TTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3719

QY 3131 AACGATCATGATTTGCTGCTATATATTCGAATATAGAGAGTACTATAGACAT 3190
Db 3718 AAATCTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3659
QY 3191 ATATTAATTAATTTCTCTCACTACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3250
Db 3658 ATACTTAATTAATTTCTCAACCAATTTCTCAATTAATTAATTAATTAATTAATTAATTAATTAAT 3599
QY 3251 GATAGATACATAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3310
Db 3598 TTCTTTTACCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3539
QY 3311 GCAATTAATTTCTTCAACCACTGAGAAACCAATTAATTAATTAATTAATTAATTAATTAATTAAT 3370
Db 3538 TACATTAATTAATTTCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3479
QY 3371 GCTGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3430
Db 3478 ACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3419
QY 3431 TAATCTCAATGCTGATTTCACTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3490
Db 3418 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3359
QY 3491 GCTGCTTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3550
Db 3358 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3299
QY 3551 GCACCATTAATGACAGAGCTTTTCAACAGAGTGGCAATTAATTAATTAATTAATTAATTAATTAAT 3609
Db 3298 TCTATCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3239
QY 3610 AACGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3669
Db 3238 TTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3179
QY 3670 AGATCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3729
Db 3178 ATTAATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3119
QY 3730 GATGCGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3789
Db 3118 TTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3059
QY 3790 TTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3849
Db 3058 CC-----CCCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3009
QY 3850 TCTTTCTGAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3909
Db 3008 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2851
QY 3910 GCAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3969
Db 2850 ACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2891
QY 3970 AACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4029
Db 2890 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2834
QY 4030 TAGTTACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4089
Db 2833 AACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2774
QY 4090 AAAAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4149
Db 2773 ATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2721
QY 4150 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4209
Db 2720 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2661
QY 4210 AAAATCACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4269

| | | | |
|----|------|-----------------------------------------------------------------|------|
| Db | 2660 | ATTATTCCTCAGCACCTAAAATTTATTTAAAATTTTAAATTAATAAATAATAATTTC | 2601 |
| Qy | 4270 | TATTTGGTAATTTTATTTATTTAAACCACTAGACAAATTAATATCCACCAATGACTATCA | 4322 |
| Db | 2600 | ACTTACGGTAAATTTCAATTCAAATCAAAATCTTCTAAACTTACTCTAAACCAATATCTCAAT | 2541 |
| Qy | 4330 | CATATGTATAGATGAAACCAATACCAACCAAGATTTATCATGTGACTCTTCACGTA | 4388 |
| Db | 2540 | AATTACTATATAAAATAAATAAACAATCTTATCTTTTAAATCTAAATCTAAACATT | 2481 |
| Qy | 4390 | TAAATCTAGATGAATTAACGACCATCTCCATGATATGAAGTCAAAATCAATCCGTAAT | 4449 |
| Db | 2480 | TAACAAAAAACAAATACCTCTTACATAAATAAAAAAAAAAAAAATTAACAAACA | 2421 |
| Qy | 4450 | TACATTTATCAACATAAT | 4468 |
| Db | 2420 | AAAAAATATCGTTAAATTT | 2402 |

RESULT 40
US-10-311

```

? Sequence 2214, Application US/10311455
? Publication No. US20030143606a1
? GENERAL INFORMATION:
? APPLICANT: OLEK, Alexander
? APPLICANT: PIEPENROCK, Christian
? APPLICANT: BERLIN, Kurt
? TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
? FILE REFERENCE: 5013.1014
? CURRENT APPLICATION NUMBER: US/10/311.455
? PRIOR FILING DATE: 2002-12-16
? PRIOR APPLICATION NUMBER: PCT/EP01/07537
? PRIOR FILING DATE: 2001-07-02
? PRIOR APPLICATION NUMBER: DE 10032529.7
? PRIOR FILING DATE: 2000-06-30
? PRIOR APPLICATION NUMBER: DE 10043626.1
? PRIOR FILING DATE: 2000-09-01
? NUMBER OF SEQ ID NOS: 2424
? SEQ ID NO 2214
? LENGTH: 11691
? TYPE: DNA
? ORGANISM: Artificial Sequence
? FEATURES:
? OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
? US-10-311-455-2214

```

| | | | | |
|-----------------------|-------|--------------------------------------------------------------|-------|-------------------------------------|
| Query Match | 2.9% | Score 135.4 | DB 14 | Length 11691 |
| Best Local Similarity | 44.6% | Seed No. 1,36-11 | | |
| Matches | 972 | Conservative | 0 | Mismatches 1176; Indels 31; Gaps 10 |
| QY | 242 | TATGCAGAACTCTTAAACTTTAAGSTCTTTACATAGAAACACTGTAATTCAGG | 301 | |
| Db | 3907 | TATCTATATACCTCAATAACGAAAAATTAATCATATATTTACTATACATATCTATAT | 3845 | |
| QY | 302 | AATACCTCTCTCTGTTGCCATTCATATTTCCATATTTAAACATATAGAGTCTT | 361 | |
| Db | 3847 | AACAACAGACATCTTTAAGCATTAATCATATACCGATTATTAAGCTTCAAAAAATTTAT | 3788 | |
| QY | 362 | TATTCGACATCATTTTAAATCAATCAAAAGAAATATAGTACTATTAATTAATA | 421 | |
| Db | 3787 | CBAATTAATAAATAAATAAACCAGAACAAATTTCCATCCCTTTAAATTAATAACATACA | 3728 | |
| QY | 422 | AAATTCCTCAATAGCTTTAGCAAGAACTACAAATATATAGACACTTAATTTAAT | 481 | |
| Db | 3727 | CTATAATTCACACAACAACCTTAACCATTAATTTAAATCCAAAAATCTATTAATTAAT | 3668 | |
| QY | 482 | AATATTCGAATTTATTAATAGCAATCATCATTTTATGATTAATTC---ACATTCAGAG | 538 | |
| Db | 3667 | AAATCTCAAAATATCACTACAAAAAATTTTAAATACAAAAATTTCTATTAATAAATACAT | 3608 | |
| QY | 539 | ATTAAACAATAGTAACACATTTCTTAACACATGATATATATCAATTAATAATTAATTT | 598 | |

| | | | |
|----|------|-----------------------------------------------------------------|------|
| Db | 3607 | AAATATTAATTAATAAACAATTAATACCTTAATTAATTAACAAATAATTAATTCATAAAATTT | 3548 |
| Oy | 599 | TTGACGTGTTCTTGATATAAATTTAGGTAAAAAAATGAAAAATCAATATTTTAAAGTGCAGAA | 658 |
| Db | 3547 | AATCTTAATATTTCTCTCTCAAAAAATTTCTATAAATAATTAATCACTACTTCATTATTA | 3488 |
| Oy | 659 | TCATTGSGCAGGAATAGCAATTCGACGTGACGACATGATATATCACTAATCTTAATG | 718 |
| Db | 3487 | ACGATTTTAAATTCATTAATAATTAATTTAAATAAATAATTTATTTATTAATTAATAA | 3428 |
| Oy | 719 | GGACGTGATCATAGAAATTTACATATAAACACTTACGAAAAAATTTAAATATTTA | 778 |
| Db | 3427 | TTATATTAATTAACCGTAATACCTAAATCCAAATCTCTTTAAAAATCTCATACATAT | 3368 |
| Oy | 779 | CCGACATATCATCTACAGCAACATATATAGCAACGACGACATATAGAGAGATTTATCT | 838 |
| Db | 3367 | CTATAACAAATTCATATTAATTTTAAATAATATATAAATACTTTATTAACCATTTAA | 3308 |
| Oy | 839 | GA--CAACAATACATGTAATTTCTTAATCTTCTCAACAACAATTAACGACATATGTA | 896 |
| Db | 3307 | AAATCAATTTAAACCATCATATTAATTAATTAATTTAAATTTATTTACTTAATTTCTAA | 3248 |
| Oy | 897 | CACAAATACACACAAACAGATATGTAATAATCTCCAAATTTATTTATTTGATGTAAG | 956 |
| Db | 3247 | AACATTTTAAATTTTAAATAAATAATTAATTAATTAATTAATTAATTAATTAATTAAT | 3188 |
| Oy | 957 | GTATCGAATTTACATCACTCAATTTCCAGTGAATAGCAGAAATTTTACATATA | 1016 |
| Db | 3187 | CATTAATAAATACATATTAACAAATATCAAAATATTAATTAATTAATTAATTAATTAAT | 3128 |
| Oy | 1017 | ATCCAAATTTTATCTCTCTGCTGCTCAAGTTTGTGTAAGAATTTGACGAACATCTTCTT | 1076 |
| Db | 3127 | AACATATTTATTAATTAATAACACATCTA-----CAACGAATTTATCACTTTACTTT | 3075 |
| Oy | 1077 | GTAATATAGTACTGCTGTGATATAAAGAACTTGATCAATTGATGAACAATTA | 1136 |
| Db | 3074 | ATATATACATACACACACCTACATCAAAATTTACTCATTAACAATTTAACTAAATAA | 3015 |
| Oy | 1137 | GTAATGATGTTCTAAATAATTAATACATCAAGATATTAAGGTGACAGTTGGACCTG | 1196 |
| Db | 3014 | TTATATCTTTAATACAAAAAATAAATAAATAAATTTAACTTAATATAATTAATCAAT | 2955 |
| Oy | 1197 | CATTAAAGAACATTAATTCGACATTCGATGGAATATCGCTGGAAAAATTTAGTATA | 1256 |
| Db | 2954 | ATATTAATAAATAAATAAATAAATAATCATCAATTAATTAATTAATTAATTAATTAAT | 2895 |
| Oy | 1257 | ATGCTGTGCTACTGTTATTAATCAAAAGGATGATGAAATTAACACATATATATACAA | 1316 |
| Db | 2894 | AAAAAATTAATTTCTTAATATATACATACAAATTAATTAATTAATTAATTAATTAAT | 2835 |
| Oy | 1317 | AAATCATTTCAATAGATCTTCATACATTAATTTGTCATCATGCTTTGGAGATCTAAT | 1376 |
| Db | 2834 | CTACAAATTAACAAATCTTTATTCAAATTTTCTATCTAAACATCACTCAATTAATCT | 2775 |
| Oy | 1377 | TAACATGTACACAGAACATTCACGCAATATGCTTTGACATAGAAATAGAAAAAGTA | 1436 |
| Db | 2774 | CTAAATTTCTAAATCATATTAATCTTAATAAATAAATAACATCATTAATTAATAAATTC | 2715 |
| Oy | 1437 | GAATATTAATTAATTAATGATCAACAACCTAAAGCTTTGTAAGGAACCATATCTGTA | 1486 |
| Db | 2714 | TACAACAATTTAAATAATATATATATATATATATATATATATATATATATATATAT | 2655 |
| Oy | 1497 | AACCATGATATCATTAAGCAAAAAAATTAATATAGAACATCAATTTGCATATCTATTT | 1556 |
| Db | 2654 | TCTTACT-----TATATATATAATTTCTAAATCCCATATTTTAAAAATCTAATAATC | 2602 |
| Oy | 1557 | ACAATTTATTAATGAAATATATTCATGATTAAGCATATCAAAATATTTATCTTAAT | 1616 |
| Db | 2601 | ACATTTTATTAATTTCAAAACCAACATATCTAAATTTCAAAAATCTCAGATCGTATTT | 2542 |
| Oy | 1617 | CTATCATCTTTGATTAATTTCAACACATGATTAATTAACGCTCAAAAAATATATGAATACA | 1676 |

Db 2541 TAATAACTTACTATTAACCTTACTATTAATCAACAACACCACTTAAATCTTCACACCA 2482
QY 1677 ACGTTCACCAATATATATTAAGACAGTCAATACCCATCGTTACATCCCACTAGTAG 1736
Db 2481 AATTT--TCTAATTTCTCATTAATAAACTCAAAATATCACTTAACATCTTA--AAATCCT 2427
QY 1737 TAACATTTTACATACAACTTAAGTATTAAGTATTAATTTCCTAGCTAATATATTAACA 1796
Db 2426 TAAATTAATAAATAAATAAATAAACAACATTCAAAATATTCATTAATATACCACTCAAA 2367
QY 1797 AATTCCTATCTCCATCTCATTAATTAACCAATGTTATCAGATTAATTAATAGTCAATTA 1856
Db 2366 AAAAAAATTCATTAATTTTAAATAAACAATTATTAACATTAATTAATTAATTAATTA 2307
QY 1857 ACTTTACTATTTATTAATAATATAGTTTGACACCTTTTTCACCTACATCATGATATA 1916
Db 2306 TTTTAACAAATTCATTTAATTCATACATTTTTCGATTTTAAATAATTTTAAACATC 2247
QY 1917 TGAACATTAATTTTACATTAATAATTAACATTAATAAAGAACCAATCATATATTAATTA 1976
Db 2246 CTCATTAATTAATTTATTAATTTCTATTAATAAATACTTATATATCTAATTTAAAAAA 2187
QY 1977 AAACACCTATTACTAGTAAATATTAATTAATGATGCTGACATTAATTAATTTATCATCC 2036
Db 2186 ATAAAAAATTTTACATTAATAATCT--CATATTAATAATACCAAAAACTCTATATACC 2128
QY 2037 TTGACATTTTCAACATTAATAATATGATACTG--TTTTTTATGCCATTAAGTCTTA 2094
Db 2127 TAAAAAACCTTTCAAAACATTAATGATTTTAAATATATTTCAATCAACATTAATAATTA 2068
QY 2095 ATAAAAATGCTAATTAACCAAAAAAGTTAAATATACATTCCTGCAACAATATGAAGAAG 2154
Db 2067 AATTAATAATTTAAAAAATAAATAAATAAATTAACATTAATCAATATTCACCAATATT 2008
QY 2155 ATACATTAACGCTAAATTAATTAACAAATTTCTAATATTAACCACTACTTAACATA 2214
Db 2007 TCTTAATCTACCTTAATTAACATTAATTAATTAATTAATTAATTAATTAATTAAT 1948
QY 2215 GGAACCAATAGTATATCATCTAGTAAAGATATCCCTACAGCATGATCAATATATAA 2274
Db 1947 AACCATTAATTAATTAATAATTAATTAATAATCAATTAATTAACA--CAAACTATTA 1889
QY 2275 GAAAAATCGTATATGATATTT--ATACAATTAACAATAATTTCACTTCAACAACAAC 2331
Db 1888 TTTTATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1829
QY 2332 TCAAGTTCTATTAATAACAATTAAGATTAATTAATTAATTAATTAATTAATTAATTA 2391
Db 1828 AACTCAACTTAATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1769
QY 2392 AATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2410
Db 1768 ACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1750

Search completed: April 6, 2004, 02:08:18
Job time : 1628 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2004, 12:18:35 / Search time 1150 Seconds
(without alignments)
12721.557 Million cell updates/sec

Title: US-10-081-051-52
Perfect score: 4750
Sequence: 1 gacccgcgaactacttag.....ccacatgacatgacac 4750

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

EST:
1: em_estba:*
2: em_estum:*
3: em_estin:*
4: em_estov:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_ges_hum:*
18: em_ges_inv:*
19: em_ges_pin:*
20: em_ges_vrt:*
21: em_ges_fun:*
22: em_ges_mam:*
23: em_ges_mus:*
24: em_ges_pro:*
25: em_ges_rtd:*
26: em_ges_pig:*
27: em_ges_vrl:*
28: gb_ges1:*
29: gb_ges2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-----------------------------|
| 1 | 157.2 | 3.3 | 1626 | 14 | CF238805 AGENCOURT |
| 2 | 151.4 | 3.2 | 1896 | 29 | CG753083 CG753083 P048-1-CO |
| 3 | 146.8 | 3.1 | 1531 | 29 | CG748014 P041-4-B0 |
| 4 | 145.2 | 3.1 | 1811 | 29 | CG753732 P048-4-G0 |

| | | | | | |
|------|-------|-----|------|----|--------------------|
| C 5 | 145 | 3.1 | 1392 | 29 | CG757503 P052-4-CO |
| C 6 | 142 | 3.0 | 1566 | 29 | CG757757 P053-1-D0 |
| C 7 | 138.6 | 2.9 | 1201 | 5 | AL565455 AL565455 |
| C 8 | 138.6 | 2.9 | 1491 | 29 | CG753221 P048-2-A0 |
| C 9 | 136.6 | 2.9 | 1348 | 29 | CG749499 P043-4-A0 |
| C 10 | 131 | 2.8 | 1376 | 13 | CG747831 P041-3-B0 |
| C 11 | 130.8 | 2.8 | 1063 | 13 | EX414736 BX414736 |
| C 12 | 127 | 2.7 | 1324 | 29 | CG746828 P040-1-G1 |
| C 13 | 125.8 | 2.7 | 1298 | 29 | CG756607 P051-4-C0 |
| C 14 | 125.8 | 2.6 | 1210 | 29 | CG749728 P044-1-C0 |
| C 15 | 125.2 | 2.6 | 1278 | 29 | CG754010 P049-2-C0 |
| C 16 | 125.2 | 2.6 | 1288 | 29 | CG744915 P037-3-F0 |
| C 17 | 124.8 | 2.6 | 1313 | 29 | CG751144 P045-3-E0 |
| C 18 | 124.8 | 2.6 | 1434 | 28 | CG748768 CH261-98P |
| C 19 | 124.4 | 2.6 | 1361 | 29 | CG744327 P036-4-E0 |
| C 20 | 124.2 | 2.6 | 1373 | 29 | CG750869 P045-2-E1 |
| C 21 | 123.4 | 2.6 | 546 | 13 | BQ455897 K624601.Y |
| C 22 | 123 | 2.6 | 1297 | 29 | CG758143 P053-3-B1 |
| C 23 | 122.8 | 2.6 | 1211 | 29 | CG747324 P040-4-D1 |
| C 24 | 121.8 | 2.6 | 1272 | 28 | CC264939 CH261-19L |
| C 25 | 121.8 | 2.6 | 1528 | 29 | CG753854 P049-1-D0 |
| C 26 | 120.6 | 2.5 | 1353 | 29 | CG744812 P037-3-B0 |
| C 27 | 120.4 | 2.5 | 1277 | 28 | CC253231 CH261-180 |
| C 28 | 120.2 | 2.5 | 1433 | 29 | CG745119 P037-4-G0 |
| C 29 | 119.6 | 2.5 | 1592 | 29 | CG750135 P044-3-D0 |
| C 30 | 119.4 | 2.5 | 1074 | 28 | BZ696936 SP_Ba09 |
| C 31 | 119.4 | 2.5 | 1101 | 29 | AL069706 DROSOPH11 |
| C 32 | 119.4 | 2.5 | 1377 | 29 | CG749971 P044-2-E0 |
| C 33 | 117.8 | 2.5 | 1124 | 13 | EX436282 BX436282 |
| C 34 | 117.8 | 2.5 | 1426 | 28 | CC231597 CH261-36A |
| C 35 | 117.4 | 2.5 | 1135 | 29 | CC233489 CH261-62N |
| C 36 | 117 | 2.5 | 1169 | 29 | AL226115 TETRADON |
| C 37 | 117 | 2.5 | 1269 | 29 | CG757211 P052-2-G0 |
| C 38 | 116 | 2.4 | 1592 | 29 | CG750135 P044-3-D0 |
| C 39 | 115.8 | 2.4 | 1061 | 13 | BX437039 BX437039 |
| C 40 | 115.6 | 2.4 | 1305 | 29 | CG744200 P036-3-H0 |
| C 41 | 115 | 2.4 | 1104 | 14 | CF264382 AGENCOURT |
| C 42 | 115 | 2.4 | 1200 | 13 | CG744200 P036-3-H0 |
| C 43 | 115 | 2.4 | 1372 | 29 | CG746616 P039-4-F0 |
| C 44 | 114.8 | 2.4 | 1454 | 29 | CG747614 P041-2-A0 |
| C 45 | 114 | 2.4 | 1296 | 29 | CG744840 P037-3-C0 |

ALIGNMENTS

RESULT 1
LOCUS CF238805
DEFINITION AGENCOURT 15099447 NICHD XGC Emb6 Silurana tropicalis cDNA clone IMAGE:6995950 5', mRNA sequence.
ACCESSION CF238805
VERSION CF238805.1 GI:33442013
KEYWORDS EST.
SOURCE Silurana tropicalis (western clawed frog)
ORGANISM Silurana tropicalis
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae; Xenopodinae; Silurana.
1 (bases 1 to 1626)
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Daniela S. Gerhard, Ph.D.
COMMENT Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabs-remail.nih.gov
Tissue Procurement: Robert M. Grainger
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be

COMMENT

Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
Class: BAC ends

FEATURES

Location/Qualifiers
1..1896

/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="mpa EcORI BAC library"
/notes="The library was generated by a partial digest of
the genomic DNA with EcORI and cloning into the BAC
vector."

ORIGIN

Query Match 3.2%; Score 151.4; DB 29; Length 1896;
Best Local Similarity 42.9%; Pred. No. 1.6e-13;
Matches 760; Conservative 0; Mismatches 1000; Indels 11; Gaps 6;

777 TACCCACATATCATCTACGAGACAACTATATGCAACGACGAAATAGACGAGTATTAT 836
1883 TAAAAAATATATATATATATATATATATATATATATATATATATATATATATAT 1824
837 CTGACACAGATACATATATATATATATATATATATATATATATATATATATATAT 896
1823 AAAAAAATATATATATATATATATATATATATATATATATATATATATATATAT 1764
897 CACAAATACAGACAAACGACATATGTATATATATATATATATATATATATATATAT 956
1763 TAAATATATATATATATATATATATATATATATATATATATATATATATATATAT 1704
957 GTATGGAATTTACATCATCTCAATTTCCAGTGAATATAGCAGAAATTTTACAAATTA 1016
1703 TAAATATATATATATATATATATATATATATATATATATATATATATATATATAT 1644
1017 ATCCATTTTATCTCTGCTGCTGCAAGTTTGTAAAGAAATGCAAGACATCTTCCTT 1076
1643 AT-TANNNAAAAAATATATATATATATATATATATATATATATATATATATATAT 1585
1077 GTATATATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1136
1584 NATATATATATATATATATATATATATATATATATATATATATATATATATATAT 1525
1137 GTATATATATATATATATATATATATATATATATATATATATATATATATATATAT 1196
1524 AATATATATATATATATATATATATATATATATATATATATATATATATATATAT 1469
1197 CATTAAGAACATATATGCAATGCTGGAATATGCTGGAATATGCTGGAATATGCTGGA 1256
1468 TATATATATATATATATATATATATATATATATATATATATATATATATATATAT 1409
1257 ATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1316
1408 TTTATATATATATATATATATATATATATATATATATATATATATATATATATATAT 1349
1317 AAAATCATTCATATAGATCTTCATATATATATATATATATATATATATATATATAT 1376
1348 TATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 1289
1377 TAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 1436
1288 TAAAAAATATATATATATATATATATATATATATATATATATATATATATATATAT 1229
1437 GAA-ATATATATATATATATATATATATATATATATATATATATATATATATATAT 1495
1228 TATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 1169
1496 AAACCACTATATATATATATATATATATATATATATATATATATATATATATATAT 1555

1168 TATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 1109
1556 TACATTTTATATATATATATATATATATATATATATATATATATATATATATATAT 1615
1108 TAAAAAATATATATATATATATATATATATATATATATATATATATATATATATAT 1049
1616 TCTATCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1672
1048 TATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 969
1673 TACACCTGTCACACATATATATATATATATATATATATATATATATATATATATAT 1732
988 AATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 929
1733 TATGTAACATATATATATATATATATATATATATATATATATATATATATATATAT 1792
928 AAAAAAATATATATATATATATATATATATATATATATATATATATATATATATAT 869
1793 AACCAATTCCTATCTCTATCTCAATTAACCAATTCCTATCTCAATTAACCAATTCCTA 1852
868 AAAAAAATATATATATATATATATATATATATATATATATATATATATATATATAT 809
1853 TTAACCTTTTATATATATATATATATATATATATATATATATATATATATATATAT 1912
808 ATAAAAATATATATATATATATATATATATATATATATATATATATATATATATAT 749
1913 ATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 1972
748 TAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 689
1973 ACTAAAAACCTTATATATATATATATATATATATATATATATATATATATATATAT 2032
688 TTTTATATATATATATATATATATATATATATATATATATATATATATATATATAT 629
2033 ATCCCTGACATATTCATATATATATATATATATATATATATATATATATATATATAT 2092
628 AATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 570
2093 TAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2152
569 ATTATATATATATATATATATATATATATATATATATATATATATATATATATATAT 510
2153 AGATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2212
509 TATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 450
2213 TAGAAACCAATATATATATATATATATATATATATATATATATATATATATATATAT 2272
449 TTTTATATATATATATATATATATATATATATATATATATATATATATATATATATAT 390
2273 AAGAAAAATGCTATATATATATATATATATATATATATATATATATATATATATATAT 2332
389 AATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 330
2333 CAAGATTCAT-TATATATATATATATATATATATATATATATATATATATATATATAT 2391
329 TATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 270
2392 AATCTATATATATATATATATATATATATATATATATATATATATATATATATATAT 2451
269 TATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 210
2452 TTATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2511
209 AATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 150
2512 TAAATGCTGAACGTTCCATTCACCAATTCACCAATTCACCAATTCACCAATTCACCA 2542
149 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 150

RESULT 3

| LOCUS | CG748014 | 1531 bp | DNA | linear | GS5 24-OCT-2003 |
|---------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------|------|--------|-----------------|
| DEFINITION | P041-4-B02.2a Ppa EcORI BAC Library Pristionchus pacificus genomic. | | | | |
| ACCESSION | CG748014 | | | | |
| VERSION | CG748014.1 GI:37968940 | | | | |
| KEYWORDS | GS5. | | | | |
| SOURCE | Pristionchus pacificus | | | | |
| ORGANISM | Pristionchus pacificus | | | | |
| REFERENCE | Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida; Nematodermatida; Pristionchus. | | | | |
| AUTHORS | Srinivasan,J., Sins,W., Jeesse,T., Wiggers-Pereboole,L., Jansen,K., Buntjer,J., van der Meulen,M. and Sommer,R.J. | | | | |
| TITLE | An integrated physical and genetic map of the nematode Pristionchus pacificus | | | | |
| JOURNAL | Mol. Genet. Genomics 269 (5), 715-722 (2003) | | | | |
| MEDLINE | 22835951 | | | | |
| PUBMED | 12984007 | | | | |
| COMMENT | Contact: Sommer RJ Evolutionary Biology Max-Planck-Institute for Developmental Biology Spemannstr. 37-39, Tuebingen D-72076, Germany Tel: 00497071601371 Fax: 00497071601498 Email: ralf.sommer@uebingen.mpg.de Class: BAC ends | | | | |
| FEATURES | Location/Qualifiers | | | | |
| Source | 1..1531 | | | | |
| | /organism="Pristionchus pacificus" | | | | |
| | /mol_type="genomic DNA" | | | | |
| | /strain="California" | | | | |
| | /db_xref="taxon:54126" | | | | |
| | /clone_1ib="Ppa EcORI BAC library" | | | | |
| | /note="The library was generated by a partial digest of the genomic DNA with EcORI and cloning into the BAC vector." | | | | |
| ORIGIN | | | | | |
| Query Match | 3.1%; Score 146.8; DB 29; Length 1531; | | | | |
| Best Local Similarity | 37.5%; Pred. No. 8.5e-13; | | | | |
| Matches 555; Conservative | 0; Mismatches 924; Indels 1; Gaps 1. | | | | |
| QY | 972 | CACCTAAATTTCCCGCTGTAAGACAGAGAAATTTTACATATATATCCATTTTATAC | 1033 | TATATC | |
| DB | 50 | CAGTTCCTCAATATCCAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN | 109 | | |
| QY | 1032 | TCTCTGATCCAGATTTCCTAAAGAAATTGCAGAACATCTTCTCTGTAGTACTTG | 1091 | | |
| DB | 110 | NNNNNNNNCCNNNNNAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN | 169 | | |
| QY | 1092 | CTGTGATTAATAAGAACTGTGTGAATCATGTATGAACAATACTATGATGTTCTTA | 1155 | | |
| DB | 170 | CCNNNAANNAANNNANNNANNNANNNANNNANNNANNNANNNANNNANNNANNNANNA | 229 | | |
| QY | 1152 | AAATATATATACCATCAGATATTTATAGGTGTACAGATTGAGCTGCATTAAGAACATPA | 1211 | | |
| DB | 230 | AAAAAANNAANNNANNNANNNANNNANNNANNNANNNANNNANNNANNNANNNANNA | 289 | | |
| QY | 1212 | TTGCATTCGATCGGAAATTAATCGGTGAAAAAATTTAGGTAATATGCTGTGTACTG | 1271 | | |
| DB | 290 | NNAANNAANNNANNNANNNANNNANNNANNNANNNANNNANNNANNNANNNANNA | 349 | | |
| QY | 1272 | TTATTAATTAAGCATGGAATGAATTAATAACCTATATATAGCAAAAAATCATTCATATG | 1331 | | |
| DB | 350 | AAAAAANNAANNNANNNANNNANNNANNNANNNANNNANNNANNNANNNANNNANNA | 409 | | |
| QY | 1332 | ATCTTCATACATTAATTTGTGCATCATGTCTTGAGACTATATTAATCATGTACACAG | 1391 | | |
| DB | 410 | AAAAAANNAANNNANNNANNNANNNANNNANNNANNNANNNANNNANNNANNNANNA | 469 | | |
| QY | 1392 | AACATTCACGCAATATGCGTTTGGACTAGAAATAGAAAAGGTAGAAATATTAATACAT | 1451 | | |

[illegible]

LOCUS CG753732 1811 bp DNA linear GSS 24-OCT-2003
DEFINITION P048-4-G03.Ya Ppa EcORI BAC Library Pristionchus pacificus genomic.
ACCESSION CG753732
VERSION CG753732.1 GI:37978509
KEYWORDS GSS.
SOURCE Pristionchus pacificus
ORGANISM Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.
REFERENCE 1 (bases 1 to 1811)
Stinvaesen,J., Sinz,W., Jesse,T., Wiggers-Perepoite,L., Jansen,K.,
Buitlcr,J., van der Meulen,W. and Sommer,R.J.
An integrated physical and genetic map of the nematode Pristionchus
pacificus
Mol. Genet. Genomics 269 (5), 715-722 (2003)
JOURNAL MEDLINE
PUBMED 22835951
COMMENT Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
Class: BAC ends.
Location/Qualifiers
FEATURES
source
1. 1811
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Ppa EcORI BAC library"
/note="The library was generated by a partial digest of
the genomic DNA with EcORI and cloning into the BAC
vector."
ORIGIN
Query Match 3.1%; Score 145.2; DB 29; Length 1811;
Best Local Similarity 35.8%; Pred. No. 1.4e-12;
Matches 577; Conservative 0; Mismatches 1030; Indels 3; Gaps 1;
726 ATCATGAAATATTTACATATTAACACCTTCCGAAAAATTAAATTTTACCCACAT 785
1760 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 1701
786 ATATCTACCGACACATATATGCAACGACATTTAGACGAAGTATATCTGACACA 845
1700 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 1641
846 ATACATGATTTATCTTAATCTTCTTACACCAATTCGACCACTATGTACACAAATAC 905
1640 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 1581
906 AACAACAAACGACATATGTGTAATAATCTCCATATTAATTTGTAGTAAGGATACGAA 965
1580 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 1521
966 TTACATCAGCAATTTCCAGTGAATAGCAGAGAAATTTTACATATTAATCCAAATT 1025
1520 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 1461
1026 TTATATCTCTGTGTCAGTTTGTCTAAAGAAATTCAGACATCTTCTTGTAGTAG 1085
1460 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 1401
1086 TACTTCTGTGATTAATAAGAACTGTGAATCTATGTAGTAAGAACTAATAGTAG 1145
1400 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 1344
1146 TTCTAATAATATATATATCAATCAAGATATATATGTGTGTACAGATGGAGCTGATTAAGA 1205
1343 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 1284

1206 ACATATATTCGAATTCATGTGGAATATCGCTGGAAAAATTTAGGTAAATATGCTGTG 1265
1283 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 1224
1266 CTACTGTTATTAATAAGGATGATGAATTAACACTATATATATGCAAAATATCAT 1325
1223 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 1164
1226 CAATAGATCTTCATACATTAATTTGTCATCATGTCTTGAGAGATCTAATTAATGTA 1385
1163 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 1104
1386 CAACGACATTCACGCAATATGCTTTTGACTAGAAATAGGAAAGGTAGAAATATA 1445
1103 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 1044
1446 ATACATTAATAGATCAACACTTAAGCTTGTGAAGAACAGTATCTGTAACCACTGA 1505
1043 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 984
1506 TATCATTTAGCAAAAACTTAATGTAGAACTACCAATTTGCATATCTATTACAAATTAT 1565
983 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 924
1566 TACATGGAATATATCTGATGATTAAGCCATATCAACATATATCTTATCTATCATCC 1625
923 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 864
1526 TTGTATATTTTCAACATATGATTAATAACGCTCAAAATATAGATATCAAGCTTCACC 1685
863 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 804
1586 ACATATTTAATAAGCACTCAATACCATGTTTACATCACTAGATATGTAACATATT 1745
803 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 744
1746 ACATACAAACTTAATGATTAATTAATTTCTAGCTAATTAATTAATCAATTCCTAT 1805
743 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 684
1806 CTCTATCTCAATTAACCAATTTGTTATCAGATTAATTAATGTCAAATTAACCTTTACT 1865
683 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 624
1866 TATTTAATAATAATATGTTGCACTTTTTCATCATCATGATATATATGACATAA 1925
623 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 564
1926 TTTTACATAAATACTAATTAATAAGAACCAATACATATTAATTAATTAACACCTTA 1985
563 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 504
1986 TTATCTAGTAATTAATTAATGATGTCGTGACATTAATTAATTAATCATCTTGAATAT 2045
503 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 444
2046 TTCAACATTAATATGATTAATCTGTTTATATGCCCATTAGTCTTAAATAAATTGCT 2105
443 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 384
2106 AATTACCCAAAAAGTTAAATACATCTTCTGCAACATATGAGAGATTAATATACAC 2165
383 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 324
2166 GTCAATTAATTAATTAATTTCTAATATATATACCACTACTCTAAGCATAGAAACCAATAG 2225
323 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 264
2226 TATATCTACTTAAGTATGATTCCTGACGATGATATCAATTAATTAAGAAATATCGTT 2285
263 AAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAAAANAAAA 204

QY 2286 ATGTATATTATACAAATTTCACCTTCAACAACAAACTCAA 2335

Dib 203 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 154

RESULT 5
CG757503

| | | | | | |
|------------|---------------------------------------|-------------|-----|--------|-----------------|
| LOCUS | CG5757503 | 1392 bp | DNA | linear | GSS 24-OCT-2003 |
| DEFINITION | P0522-4-008. za Ppa EcORI BAC Library | | | | |
| ACCSSION | CG5757503 | | | | |
| VERSION | CG5757503 | | | | |
| KEYWORDS | Genomic survey sequence. | | | | |
| SOURCE | CG5757503.1 | GI:37986131 | | | |
| ORGANISM | GSS. | | | | |
| | <i>Pristionchus pacificus</i> | | | | |
| | <i>Pristionchus pacificus</i> | | | | |
| | <i>Pristionchus pacificus</i> | | | | |

FEATURES

```

FEATURES
    source
        location/Qualifiers
            1..1392
                /organism="Pristionchus pacificus"
                /mol_type="genomic DNA"
                /strain="Californina"
                /db_xref="taxon:54126"
                /clone_lib="Ppa EcoRI BAC library"
                /note="The library was generated by a partial digest of
                the genomic DNA with EcoRI and cloning into the BAC
                vector."

```

ORIGIN

| | | | | |
|-----------------------|-----------------|--------------------|------------|--------------|
| Query Match | 3.1%; | Score 145; | DB 29; | Length 1392; |
| Best Local Similarity | 46.5%; | Pred. No. 1.6e-12; | | |
| Matches 591; | Conservative 0; | Mismatches 658; | Indels 22; | Gaps 4 |

[illegible]

| | | | |
|----|------|----------------------------------------------------------------|------|
| QY | 1780 | AGCAATATATTAACAATATCCATATCCATCTGCATTTACCAATGGTATCAGATTA | 1833 |
| Db | 993 | AAATAATTTAAATATATATAAATAATATATAAATAAATAAATAAATAATTAATTA | 934 |
| QY | 1840 | TAAATTAAGTCAAAATTAACCTTTACTATATTTATAAAAATATATGTTTGCACATTTTCA | 1895 |
| Db | 933 | TATATAAATATATAAATAAATAAATAAATAATATATAAATAATTAATAAATAATATAA | 874 |
| QY | 1900 | CTACATCATGTATATTAAGAACAATATTTACATTAATAATACATATAAAGAACATATA | 1959 |
| Db | 873 | TTATATTTATATAAATAAATAAATAAATAATTAATAAATAATATATATTA-----ATTA | 825 |
| QY | 1960 | CATATATTAATTAATCTATAAACACCTATTACTACGTATATAATACATAGATGCTGCAT | 2019 |
| Db | 824 | AATATAATTAATAATATATATATAAATAAATAAATAAATAATATATATATAATAATAT | 765 |
| QY | 2020 | AATATAATTCATCATCTCGATGCATATTTCAAGCAATAAATATGTAATACGTTTTTTAT | 2079 |
| Db | 764 | AATTAATTTATTAATAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAATTA | 707 |
| QY | 2080 | GCCCATTAGTCTTAATTAATTAATGCTAATTAACCCAAAGGTTAAATAACATCTCTGC | 2139 |
| Db | 706 | AAAAAAAAAAAAATATAAATAATTTAAATTAATAAATAATTAATAATTAATAATATAA | 647 |
| QY | 2140 | AACATATGAAGAAGATACAAATACAGTCMAATTAATAACAATTTCTATATATACA | 2199 |
| Db | 646 | AAAAATATAAATAAATAAATAAATAATTTATTAATAAATAAATAAATAAATAATATATAA | 590 |
| QY | 2200 | CCTACTCTAAACATAGGACCAATAGTATATCATACCTAAGTAAGATATCCCTACACAT | 2259 |
| Db | 589 | ---AAATATAAATAAATAAATAAATAATTTTAAATAATTTATTAATTTATTTATTT | 533 |
| QY | 2260 | GTATCAATATATAAGAAAAATGCTATGTATATTTATACATTTACAAATTTTCCAACTT | 2319 |
| Db | 532 | TTATTAATTTATTAATTTATATATATTTATTAATAATATAAATAAATAAATAAATAATTA | 473 |
| QY | 2320 | CAACACAAAACTCA--AAGTCTATTTATACCAATTCAGAGATTTATATCTTACCC | 2376 |
| Db | 472 | AAAAAATAAATAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAATTT | 413 |
| QY | 2377 | AACATATGTTATCAATTTCTATATATCATATCTCTTTTAAAAAGATGACTTCACA | 2436 |
| Db | 412 | AAAAATATATATAAATAATATATATTTTAAAAATAAATAAATAAATAAATAAATAATATATA | 353 |
| QY | 2437 | TTCATCATTAATATATATACATGCAATGTTTATTCGAACAACATTACACATCCAAAT | 2496 |
| Db | 352 | TAAAAATAAATAAATAAATAATATATTAATTTATTTATTTATTAATAATATATAATTTAT | 293 |
| QY | 2497 | TTTCAATTTACGGTATATGCTGAACGTCATTCACAACAATTCAGACAGTGGGTGATAC | 2556 |
| Db | 292 | AAATTAATATATATATAAATAATTTATATATAATATAAATAAATAAATAATATAATTAATA | 233 |
| QY | 2557 | ATTACCTCCAAAACGAGATACAGAGTTTAAATCTGTCTTCAATCTTCAGACAGC | 2616 |
| Db | 232 | AAAAAATAAATAAATAAATAATATATTAATTTTATTAATAAATAAATAAATAAATAATTA | 173 |
| QY | 2617 | TTTACGTAATGTAACAGAAATAACTTTAGCACTAAACAATCATTAATAGCGCAATTAGA | 2676 |
| Db | 172 | TATATATTTATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAATTT | 113 |
| QY | 2677 | TACTATATCTA 2687 | |
| Db | 112 | NNNNNNCACTA 102 | |

RESULT 6

CG757757/c 1566 bp DNA linear GSS 24-OCT-2003
 LOCUS CG757757
 DEFINITION P053-1-D07.zc Ppa EcoRI BMC Library *Pristionchus pacificus* genomic.
 accession genomic survey sequence.
 CG757757
 accession CG757757.1 GI:37986636
 version
 keywords GSS.

Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 9232.f. For more information about this cluster, see

http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DF005BH09NP1&cluster=9232.f. Contact : Peng Liang Email : filiang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue Genoscope, sequence ID : CS0DF005BH09NP1.

FEATURES

Source

1. 1201

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DF005Y018"

/issue_type="FETAL BRAIN"

/dev_stage="fetal"

/clone_lib="Homo sapiens FETAL BRAIN"

/note="Organ: Brain; Vector: PCWVSPT6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the PCWVSPT6 vector. Library was not normalized."

ORIGIN

```

Query Match      2.9%; Score 138.6; DB 9; Length 1201;
Best Local Similarity 37.7%; Pred. No. 1.6e-11;
Matches 330; Conservative 139; Mismatches 403; Indels 3; Gaps 3;

QY 1437 GAAATTAATACATTAATGATGACACAACTTAAGTGTGAGAGAACCTACTGTAA 1496
DB 318 GAAATTAATACATTAATGATGACACAACTTAAGTGTGAGAGAACCTACTGTAA 1496
QY 1497 AACCACTGATATGATGACACAACTTAAGTGTGAGAGAACCTACTGTAA 1556
DB 378 AGCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 437
QY 1557 ACAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1616
DB 438 ATATATATATATATATATATATATATATATATATATATATATATATAT 497
QY 1617 CTATCATCTCTGATATATATATATATATATATATATATATATATATATAT 1675
DB 498 ATATATATATATATATATATATATATATATATATATATATATATATAT 557
QY 1676 AAGCTTCAACCATATATATATATATATATATATATATATATATATAT 1735
DB 558 ATATATATATATATATATATATATATATATATATATATATATATATAT 617
QY 1736 GTATCATATATATATATATATATATATATATATATATATATATATATAT 1795
DB 618 ATATATATATATATATATATATATATATATATATATATATATATATAT 677
QY 1796 AAATTCCTATCT-CTATCTCAATTAACCAATGTTATGAGATTATAAATAGCAATT 1854
DB 678 ATATATATATATATATATATATATATATATATATATATATATATATAT 737
QY 1855 AAATTCCTATCTATATATATATATATATATATATATATATATATATAT 1914
DB 738 ATATATATATATATATATATATATATATATATATATATATATATATAT 797
QY 1915 AATGACATATATATATATATATATATATATATATATATATATATATATAT 1974
DB 798 ATATATATATATATATATATATATATATATATATATATATATATATAT 857
QY 1975 TAAATACACCTATATATATATATATATATATATATATATATATATATAT 2033
DB 858 ATATATATATATATATATATATATATATATATATATATATATATATAT 917
QY 2034 TCCTTGACATATATATATATATATATATATATATATATATATATATATATAT 2093
DB 918 ATATATATATATATATATATATATATATATATATATATATATATATAT 977
QY 2094 AATTAATATGCAATATATATATATATATATATATATATATATATATATAT 2153

```

```

DB 978 ATATATATATATATATATATATATATATATATATATATATATATATAT 1037
QY 2154 GATCAATATACACGCAATATATATATATATATATATATATATATATATATATAT 2213
DB 1038 ATATATATATATATATATATATATATATATATATATATATATATATAT 1097
QY 2214 AGAACAATATATATATATATATATATATATATATATATATATATATATAT 2273
DB 1098 ATATATATATATATATATATATATATATATATATATATATATATATAT 1157
QY 2274 AGAATAATCGTTATATATATATATATATATATATATATATATATATATAT 2308
DB 1158 ATATATATATATATATATATATATATATATATATATATATATATATAT 1192

```

```

RESULT 8
CG753221/c 1491 bp DNA linear GSS 24-OCT-2003
LOCUS P048-2-A01.2a Ppa EcoRI BAC Library Pristionchus pacificus genomic,
DEFINITION CG753221
ACCESSION CG753221.1 GI:37977480
VERSION GSS.
KEYWORDS Pristionchus pacificus
SOURCE Pristionchus pacificus
ORGANISM Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Mediolagasteridae; Pristionchus.
1 (bases 1 to 1491)
AUTHORS Striavasan, V., Sim, W., Jesse, T., Wiggers-Perebolte, L., Jansen, K.,
Buntjer, J., van der Meulen, M., and Sommer, R.J.
TITLE An integrated physical and genetic map of the nematode Pristionchus
pacificus
MOL. GENET. GENOMICS 269 (5), 715-722 (2003)
JOURNAL MEDLINE
PUBMED 12884007
COMMENT Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
Class: BAC ends
Location/Qualifiers
1. 1491
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Ppa EcoRI BAC Library"
/note="The library was generated by a partial digest of
the genomic DNA with EcoRI and cloning into the BAC
vector."

```

```

ORIGIN
Query Match      2.9%; Score 138.6; DB 29; Length 1491;
Best Local Similarity 39.2%; Pred. No. 1.5e-11;
Matches 522; Conservative 0; Mismatches 808; Indels 0; Gaps 0;

QY 1098 ATATATAAGAACTGTTGATCATGATGACAAATATAGTATGATGTTCTTAATAATTA 1157
DB 1360 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1301
QY 1158 TATATCAATCAAGATATATATATATATATATATATATATATATATATATATAT 1217
DB 1300 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1241
QY 1218 TTGCATGTGATATATATATATATATATATATATATATATATATATATATATAT 1277
DB 1240 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1181
QY 1278 CTAAAGCATGAAATATTAATAACATATATATATATATATATATATATATATATAT 1337

```

```

Db      1180 NAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1121
Qy      1338 ATACATTAATGGTCATCATGCTTGGAGATCTATATTAACATGTCACAGACATTT 1397
Db      1120 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1061
Qy      1398 CACGCAATATGGCTTTGGATGAGAAATAGGAGAAAGTGAGAAATATATATCATTAATAG 1457
Db      1060 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1001
Qy      1458 ATCAACAACCTTAAGCTTGTGAAGACAGACTGTAAACCACTGATATCATTAAGCA 1517
Db      1000 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 941
Qy      1518 AAAAACTTAATGTAAGTACCAATTTGCAATCTATTTACAAATTTATTAATGAGATA 1577
Db      940 NAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 881
Qy      1578 TATCACTAGATTAAGCCATATCAAAATATATCTTGTATCATCTTGTATATTTTC 1637
Db      880 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 821
Qy      1638 AACACATGATTAATTAACGCTCAAAATATAGAAATACACGCTTCCACATATATATAA 1697
Db      820 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 761
Qy      1698 AGCGTCAATTAAGCCATGCTTAACATCACTATGTAATGTAATTTACATTAACAACT 1757
Db      760 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 701
Qy      1758 TAAATTAATCTTAATTTCTAGCTAATTAATTAACAAATTTCTCTCTCTCA 1817
Db      700 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 641
Qy      1818 TTAACCAATTTGTTATCATATTAATTAAGTCAATTTAACTTTTACTTATTTAAAT 1877
Db      640 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 581
Qy      1878 AATATGTTGACACTTTTCTCATCATCATGATATATATGACATATTTTACATAAA 1937
Db      580 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 521
Qy      1938 TAACTTAATTAAGCAATATCATATTTAAATTAATTAACCTATTTACTAGCTAA 1997
Db      520 NAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 461
Qy      1998 TATTCATAGATGTCCTGACATTAATTAATTTCTATCATCTTGAATTTTCAACATTA 2057
Db      460 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 401
Qy      2058 ATATGATTAAGTCTTTTATTAAGCCATTAAGCTTCTTAATTAATTTGCTAATTAACCAAA 2117
Db      400 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 341
Qy      2118 AGGTTAAATTAACATCTCTGCAATTAATTAAGAAATTAAGAAATTAAGTAAAT 2177
Db      340 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 281
Qy      2178 AAAAAATTTCTAATTAATTAACACTACTTAACATAGGAACCAATAGTATATCATCT 2237
Db      280 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 221
Qy      2238 AAGTAAGTATCTCTGACATGATTAATTAAGAAATTAAGAAATTAAGTAAAT 2297
Db      220 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 161
Qy      2298 ACAAATTAACAAATTTCCAACTTCAACAACAACCAAGTCTAATTAATTAACAAATTTCAA 2357
Db      160 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 101
Qy      2358 GAATTAATTAATCTTCAACATATGATTAATTAATTAATTAATTAATTAATTAATTT 2417

```

```

Db      100 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATTTCTCATTAAT 41
Qy      2418 TAAAAAGATG 2427
Db      40 GTACTCGATG 31

RESULT 9
CG749499/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

COMMENT
Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spermanstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@uebingen.mpg.de
Class: BAC ends.

FEATURES
source
Location/Qualifiers
1..1348
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="pPa EcoRI BAC library"
/note="The library was generated by a partial digest of
the genomic DNA with EcoRI and cloning into the BAC
vector."

ORIGIN
Query Match 2.9%; Score 136.6; DB 29; Length 1348;
Best Local Similarity 46.9%; Pred.No.3.1e-11;
Matches 554; Conservative 0; Mismatches 616; Indels 12; Gaps 4;

Qy      1277 ACTAAAGCATGATGAATTAATTAACACTATATATAGCAAAAAATCATTCATAGATCTT 1336
Db      1342 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1283
Qy      1337 CATCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1396
Db      1282 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1227
Qy      1397 TCACGCAATATGCTTTGACATGAAATTAAGAAATTAAGAAATTAATTAATTAATTA 1456
Db      1226 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1167
Qy      1457 GATCAACACTTAAGCTTTGACAGAAACAGTACTGTAACCACTGATATCATTAAGCA 1516
Db      1166 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1109
Qy      1517 AAAAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1576
Db      1108 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1049
Qy      1577 ATATCACTAGATTAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTT 1636

```


ORIGIN

/clone.lib="Ppa EcoRI BAC Library"
/note="The library was generated by a partial digest of
the genomic DNA with EcoRI and cloning into the BAC
vector."

```

Query Match      2.6%; Score 125.8; DB 29; Length 1210;
Best Local Similarity 44.4%; Pred. No. 1.3e-09;
Matches 465; Conservative 0; Mismatches 575; Indels 8; Gaps 3;

1453 AATGATCAGACCTTAAGCTGTTGAGAGACCAAGTCTGTAACCACTGATATCAT 1512
1210 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1151
1513 AGCAAAAAAATTATGAGTACCAATTTGCAATCTATTACCAATTTATTCATGA 1572
1150 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1091
1573 GAATATATCACTAGATTAAGCAATCAATTAATTAATTAATTAATTAATTAAT 1632
1090 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1031
1533 ATTCAACACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1692
1030 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 971
1693 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1752
970 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 911
1753 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1812
910 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 851
1813 CTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1872
850 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 791
1873 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1932
790 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 731
1933 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1991
730 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 671
1992 ACGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 611
670 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 611
2048 CAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2107
610 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 551
2108 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2167
550 NTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 491
2168 CAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2227
490 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 431
2228 TATCATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2287
430 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 372
2288 GTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2347
371 --AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 314
2348 ACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2407
313 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 254

```

```

Query Match      2.6%; Score 125.2; DB 29; Length 1276;
Best Local Similarity 41.2%; Pred. No. 1.6e-09;
Matches 445; Conservative 0; Mismatches 634; Indels 0; Gaps 0;

1417 ACTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1476
1275 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1216
1477 TGAAGAACAGTACTGTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1536
1215 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1156
1537 ACGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1596
1155 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1096
1597 ATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1656
1095 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1036
1657 GGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1716
1035 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 976

```

FEATURES

source

Location/Qualifiers

1..1276

/organism="Pristionchus pacificus"

/mol_type="genomic DNA"

/strain="California"

/db_xref="taxon:54126"

/clone.lib="Ppa EcoRI BAC Library"

/note="The library was generated by a partial digest of
the genomic DNA with EcoRI and cloning into the BAC
vector."

ORIGIN

```

Query Match      2.6%; Score 125.2; DB 29; Length 1276;
Best Local Similarity 41.2%; Pred. No. 1.6e-09;
Matches 445; Conservative 0; Mismatches 634; Indels 0; Gaps 0;

1417 ACTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1476
1275 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1216
1477 TGAAGAACAGTACTGTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1536
1215 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1156
1537 ACGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1596
1155 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1096
1597 ATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1656
1095 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1036
1657 GGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1716
1035 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 976

```


QY 2140 AACATATGAGAGATACATACACGCTCAATTAATTAACAAATTTCTAATATACCA 2199
Db 532 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 473
QY 2200 CCTACTTAAACATAGAACCAATAGTATCTCATCTAGTAAAGTATCCCTACAGCAT 2259
Db 472 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 413
QY 2260 GTATCAATATATATAGAAAAATCGTATATATTATACAAATTAACAAATTTCCAACTT 2319
Db 412 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 353
QY 2320 CAATACCAATCTCAAGTCTATATATACAAATTCAGAGATATATATCTTACCAAC 2379
Db 352 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 293
QY 2380 ACATAGTATCAATTTCTATATATACATCATATCTTTTAAAGATGACTTACAACTT 2439
Db 292 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 233
QY 2440 ATATATATATATATATATATATATATATATATATATATATATATATATATAT 2490
Db 232 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 182

RESULT 17
CG751144/c 1313 bp DNA linear GSS 24-OCT-2003
LOCUS P045-3-B08.2a Ppa EcoRI BAC Library Pristionchus pacificus genomic,
DEFINITION genomic survey sequence.
ACCESSION CG751144
VERSION CG751144.1 GI:37973308
KEYWORDS GSS.
SOURCE Pristionchus pacificus
ORGANISM Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.
1 (bases 1 to 1313)
Buntjer,J., van der Meulen,W. and Sommer,R.U.
An integrated physical and genetic map of the nematode Pristionchus
pacificus
Mol. Genet. Genomics 269 (5), 715-722 (2003)

JOURNAL MEDLINE
PUBMED 22835951
COMMENT Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Semmerstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
Class: BAC ends.
Location/Qualifiers

FEATURES
source 1.1313
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Ppa EcoRI BAC Library"
/note="The library was generated by a partial digest of
the genomic DNA with EcoRI and cloning into the BAC
vector."

ORIGIN
Query Match 2.6%; Score 124.8; Db 29; Length 1313;
Best Match Similarity 41.8%; Pred. No. 1.9e-09;
Matches 489; Conservative 0; Mismatches 682; Indels 0; Gaps 0;

QY 1293 AATTTAAACATATATATATATATATATATATATATATATATATATATATATAT 1352
Db 1310 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1251

QY 1353 CATCATGCTTGGAGCTATATATATTAACATGTACAGACATTCACGCAATATGCTT 1412
Db 1250 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1191
QY 1413 TTGAGCTAGAAATGGAAGAGTGAAATATATATATATATATATATATATATATAT 1472
Db 1190 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1131
QY 1473 TTGTTGAGGAGACGATCTGTATTAACCACTGATATCATTAAGCAAAATTTATGTAG 1532
Db 1130 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1071
QY 1533 AACTTACCAATTTGATATCTATTTTACAAATTTATTAATGATATATCTGATATAG 1592
Db 1070 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1011
QY 1593 CCATATCAACATATATCTTATGCTATCTATCTTGTATATATTTCAACACAAATGATAAT 1652
Db 1010 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 951
QY 1653 AAACGCTCAAAATATATGAAATACACGCTTACACACATATATTAATAAGCAAGTCAATACC 1712
Db 950 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 891
QY 1713 ATCGTTAATCACTCAATCAATATATATATATATATATATATATATATATATATAT 1772
Db 890 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 831
QY 1773 ATTTCTAGCTAATATATATATATATATATATATATATATATATATATATATAT 1832
Db 830 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 771
QY 1833 CAGATATATATATATATATATATATATATATATATATATATATATATATATAT 1892
Db 770 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 711
QY 1893 TTTTCTACTCATCATATATATATATATATATATATATATATATATATATATAT 1952
Db 710 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 651
QY 1953 ACCATATATATATATATATATATATATATATATATATATATATATATATATAT 2012
Db 650 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 591
QY 2013 CTGACATATATATATATATATATATATATATATATATATATATATATATATAT 2072
Db 590 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 531
QY 2073 TTTTATGCCCATATATATATATATATATATATATATATATATATATATATAT 2132
Db 530 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 471
QY 2133 TTCTGCAACATATATATATATATATATATATATATATATATATATATATATAT 2192
Db 470 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 411
QY 2193 ATTAACCACTTATCTTAAACATAGAACCAATATATATATATATATATATATAT 2252
Db 410 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 351
QY 2253 ACAGCATATATATATATATATATATATATATATATATATATATATATATATAT 2312
Db 350 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 291
QY 2313 CCAATCTCAACCAAACTCAAGTCTATATTAACAAATTTCAAGATTTATATATCTT 2372
Db 290 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 231
QY 2373 CACCAACATATATATATATATATATATATATATATATATATATATATATAT 2432
Db 230 NAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 171
QY 2433 ACATATCATATATATATATATATATATATATATATATATATATATATATAT 2493

Db 170 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 140

RESULT 18
CC187638 1434 bp DNA linear GSS 08-MAY-2003
LOCUS CH261-98P15.Spe.1 CH261 Gallus gallus genomic clone CH261-98P15,
DEFINITION genomic survey sequence.
ACCESSION CC187638
VERSION CC187638.1 GI:30431538
KEYWORDS GSS:
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 1434)
AUTHORS Kremitzki, C., Higinbotham, J., Wylie, K., Carter, J., McPherson, D.,
Warren, W., Graves, T., Mardis, E. and Wilson, R.
Gallus gallus BAC End Reads
Unpublished (2003)
Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 18200 Std Error: 0.00
Seq primer: Spe ATTGAGTGCACCTATAG
Class: BAC ends
High quality sequence start: 422
High quality sequence stop: 487.
Location/Qualifiers
1. 1434
/organism="Gallus gallus"
/mol_type="genomic DNA"
/strain="Red Jungle Fowl"
/db_xref="taxon:9031"
/clone="CH261-98P15"
/sex="female"
/cell_line="UCD001, inbred 256"
/clone_lib="CH261"
/note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;
CH261 Female Chicken library - For library and clone
ordering information: <http://www.chori.org/bacpac>"

ORIGIN

Query Match 2.6%; Score 124.8; DB 28; Length 1434;
Best Local Similarity 40.6%; Pred. No. 1.8e-09;
Matches 567; Conservative 0; Mismatches 830; Indels 0; Gaps 0;

1098 ATATATAAAGACTGGTGAATCATGATGAAACATAGTATGATGTTCTAAATAA 1157
Db 38 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 97
1158 TATACATCAAGATATTAGGTGACATGAGCTGCATTAAAGAACATATTGCA 1217
Db 98 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 157
1218 TTGCATGTGAATTAATCGCTGGAATAATTTAGTAAATGCTGTGCTACTGTATA 1277
Db 158 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 217
1278 CTAAAGCATGATGAATTAACACTATATATAGCAAAATCATTCATAGATCTTC 1337
Db 218 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 277
1338 ATACATTATTTGCTCATCATGCTTGAGATCTATATTTACATGTCACAGAGATT 1397
Db 278 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 337
1398 CAGCAATATGCTTTTGACTAGAAATAGAAAAGTAGAATATATATACATTATAG 1457
Db 338 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 397

QY 1458 ATCAACAATTAAGCTTTGTAAGAACGAGTACTGTAACCACTGATATCATAGCA 1517
Db 398 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 457
1518 AAAAATCTTAATGTGAATCACTCAATTTGCATATCTATTTCATTTATTCATGAGATA 1577
Db 458 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 517
1578 TATCACTAGATAAGCCATATCAACATATATCTTATGTTATATCTTGTATTTTC 1637
Db 518 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 577
1638 AACCAATGATTAATTAACGTCGAAATATAGATCAAGTTCACCATATATATATAA 1697
Db 578 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 637
1698 AGCAGTCAATACCCATCGTTAATCATCACTAAGTATGTAACATATTACATACAACT 1757
Db 638 AAAAAAAAAACCCCCCAAAAAAAAAAAAAAAAAA 697
1758 TAAGTATTAATTAATTTCTGCTAATATATTAACAAATTCATCTCTATCTCA 1817
Db 698 CAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 757
1818 TTAACCAATTTGTTATCAGATTATTAATTAAGTCAATTTAACTTTATTTATTAAT 1877
Db 758 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 817
1878 AATATGTTTGCACATTTTTCATCATCTGATATTAAGACATATTTTACATTA 1937
Db 818 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 877
1938 TAACTAATTAAGAACCAATATATATTAATTTCTAATAACCACTTTTACTAGTAA 1997
Db 878 AAAAAAAAAACCAACCAAAAAAAAAAAAAAAAAA 937
1998 TATTAATGATGTCCTGATATTAATTTCTATCTCTGACATATTTCAACATTA 2057
Db 938 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 997
2058 ATATGATTAATCTGTTTTTTATGCCCCTAATGTTCTTAATTAATTTGCTAATTA 2117
Db 998 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1057
2118 AGCTTAATAATCACATCTCTGCAACATATGAAGAGATACATACGCAATTAATT 2177
Db 1058 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1117
2178 AAACAAATTTCTAATATTAACCACTCTCTAATAGAGCAATAGTATCATACCT 2237
Db 1118 AAAAAAAAAAACAACAAAAAAAAAAAAAAAAA 1177
2238 AAGTAAATATCCCTACAGATGATCAATTAAGAAAAATCGTATGTTATTTAT 2297
Db 1178 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1237
2298 ACATTAACAATTTTCAACCTTCAACCACTCAAACTCAAGTTCTATTAACAATTC 2357
Db 1238 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1297
2358 GAATTAATATCTTCAACCACTGATTTATCAATTTCTAATATACATCATATCTTTT 2417
Db 1298 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1357
2418 TAAAGATGATCTCAATTCATATCATTAATATATATCATGCAATGTTATTCAGAA 2477
Db 1358 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1417
2478 CATTAACAATTCAGAA 2494
Db 1418 AAAAAAAAAAAAAAAAAA 1434

RESULT 19
CG744327/c 1361 bp DNA linear GSS 24-OCT-2003
LOCUS P036-4-E08.ya Ppa EcorI BAC Library Pristionchus pacificus genomic,
DEFINITION genomic survey sequence.
CG744327
ACCESSION CG744327.1 GI:37965195
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
1 (bases 1 to 1361)
Strimvasan, U., Sinz, W., Jesse, T., Wiggers-Perebolte, L., Jansen, K.,
Buntjer, J., Van der Meulen, M., and Sommer, R.J.
An integrated physical and genetic map of the nematode Pristionchus
pacificus
Mol. Genet. Genomics 269 (5), 715-722 (2003)

TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT
Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel.: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
Class: BAC ends.

FEATURES
Source
Location/Qualifiers
1..1361
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/note="The library was generated by a partial digest of
the genomic DNA with EcorI and cloning into the BAC
vector."

ORIGIN

Query Match 2.68; Score 124.4; DB 29; Length 1361;
Best Local Similarity 41.95; Pred. No. 2.1e-09;
Matches 452; Conservative 0; Mismatches 627; Indels 0; Gaps 0;

1417 ACTGAAATAGGAAAGGTGAATATTAATACATTAATGATCAGACCTAAAGCTTGT 1476
1306 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1247
1477 TGAAGGACGAGTACTGTAAACCACTGATATCATTAGCAAAAACCTTAATGTAGACT 1536
1246 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1187
1537 ACCAATTGATATCTATTACAAATTATATACATGGAATATTCCTGATTAAGCCAT 1596
1186 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1127
1597 ATCAACATATTAATCTTGTCTATCATCTCTTGTATTAATTCACACATGATTAATTAAC 1656
1126 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1067
1657 GGTCAAAATATGAAATACAAAGTTCACCAATATTAATAAGCATCAATTAACCATG 1716
1066 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1007
1717 TTAACATCAACTAGATGTAACATATTAATTAACAACTTAAGTATTAATTAATTT 1776
1006 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 947
1777 CTAGCTATTAATTAACAAATTCCTATCTCTATCTCAATTAACCAATTTGTATCAGA 1836
946 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 887

QY 1837 TTATAATTAAGTCAATTAATTAATTTACTATTTAATAATTAATGTTGCAACCTTT 1896
DB 886 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 827
QY 1897 TCATACATCACTGATATTAATGAAATATTTTCACTAAATTAATTAATTAAGAACCA 1956
DB 826 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 767
QY 1957 ATACATATTAATTAATTAATTAACACCTATTACTAGTAATTAATTAATGATGCTGA 2016
DB 766 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 707
QY 2017 CATAATTAATCTATCATCTCTGATATTTCAATTTCAATTAATTAATGATTAATCTTTT 2076
DB 706 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 647
QY 2077 TATGCCCATTTAGTCTTATTAATTAATTTGCTATTTACCAAAAGTTAAATACACATTC 2136
DB 646 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 587
QY 2137 TGCAACATATGAAAGATACATACACGCAATTAATTAATTAATTTCTATATTA 2196
DB 586 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 527
QY 2197 CCACCTACTTAACATAGAACCAATAGTATATCATACCTAAGTAAAGATATCCCTACAG 2256
DB 526 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 467
QY 2257 CATGTATCAATTAATTAAGAAATCGTATGATTAATTTATACATTAACAAATTTCCAA 2316
DB 466 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 407
QY 2317 CTCAACACAGAACTCAAGTTCTATTAATTAATTAATTAATTAATTAATTAATTAATTA 2376
DB 406 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 347
QY 2377 AACCATGTTATCAATTTCTATTAATCATCATATCTTTTAAAGATGACTCAAA 2436
DB 346 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 287
QY 2437 TTCATACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2495
DB 286 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 228

RESULT 20
CG750869/c 1373 bp DNA linear GSS 24-OCT-2003
LOCUS P045-2-E10.za Ppa EcorI BAC Library Pristionchus pacificus genomic,
DEFINITION genomic survey sequence.
CG750869
ACCESSION CG750869.1 GI:37972758
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
1 (bases 1 to 1373)
Strimvasan, U., Sinz, W., Jesse, T., Wiggers-Perebolte, L., Jansen, K.,
Buntjer, J., Van der Meulen, M., and Sommer, R.J.
An integrated physical and genetic map of the nematode Pristionchus
pacificus
Mol. Genet. Genomics 269 (5), 715-722 (2003)

TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT
Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel.: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
Class: BAC ends.

FEATURES

Location/Qualifiers
 1..1373
 /organism="Pristionchus pacificus"
 /mol_type="genomic DNA"
 /strain="California"
 /db_xref="taxon:54126"
 /clone_lib="Ppa EcoRI BAC Library"
 /note="The library was generated by a partial digest of the genomic DNA with EcoRI and cloning into the BAC vector."

ORIGIN

Query Match 2.6%; Score 124.2; DB 29; Length 1373;
 Best Local Similarity 41.3%; Pred. No. 2.2e-09;
 Matches 441; Conservative 0; Mismatches 627; Indels 0; Gaps 0;

1420 AGAATATGAGAAAGTGAATATTAATCATTAATAGTCACAACTTAAGCTTGTGA 1479
 1356 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1297
 1480 AGGAACCACTAGTAAACCACTGATATCATTAACCAAACTTAAGTGAAGTACC 1539
 1296 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1237
 1540 AATTGCAATCTATTACATTTTATCATGAGATATTCACCTAGTAAAGCCATATC 1599
 1236 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1177
 1600 AATCATATATCTTGTATCATCTCTGATTAATTTCAACATGATTAATTAAGGT 1659
 1176 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1117
 1660 CAAATATAGATGACAGCTTCCACATATATTAAGACATCAATACCATCTGA 1719
 1116 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1057
 1720 ACATCAACTAGTATGATATATATTAATTAAGTAACTTAATTAATTTTCT 1779
 1056 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 997
 1780 AGCTAATATATTAACAAATCTCTATCTCAATTAACCAATTTGTATGATTA 1839
 936 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 937
 1840 TAAATATGCAATTAACCTTTTACTTATTAATTAATTAATTTGCAACTTTTCA 1899
 936 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 877
 1900 CTACATCATGATATATGAACATTAATTTTCAATTAATTAATTAATTAATTA 1959
 876 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 817
 1960 CATATATTAATTAATTAACCACTATTTACTAGCTAATTAATTAATTAATTT 2019
 816 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 757
 2020 AATAATTTCTATCATCTTGAATTTTCAACATTAATTAATTAATTTTAT 2079
 756 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 697
 2080 GCCCTTGTCTTAAATTTGCTAATTTCCCAAAAGTTAAATTAATTAATTTCT 2139
 656 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 637
 2140 AACATATGAGAAAGTGAATATTAATTAATTAATTAATTTTCAATTAATTA 2199
 636 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 577
 2200 CTTACTTAAACATGAGAACCAATGATATCATCTTAAGTAAAGATTCCTTAC 2259
 576 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 517
 2260 GTATCAATATTAAGAAATATGATATTTATTAATTAATTAATTAATTTCA 2319

Db 516 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 457
 Qy 2320 CAACCAAAACTCAAGTTCTATTATTAACAAATTCAGAAATTAATTAATTA 2379
 Db 456 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 397
 Qy 2380 ACATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2439
 Db 396 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 337
 Qy 2440 ATACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2487
 Db 336 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 289

RESULT 21
 BQ455897/c 546 bp mRNA linear EST 29-MAY-2002
 LOCUS k22401.y1 Dirofilaria immitis adult PAMP1 v1 Dirofilaria immitis
 DEFINITION cDNA 5' similar to TR:P90551 P90551 Glycerol-3-phosphate
 DEHYDROGENASE, mRNA sequence.
 ACCESSION BQ455897
 VERSION BQ455897.1 GI:21259002
 KEYWORDS EST.
 SOURCE Dirofilaria immitis
 ORGANISM Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioides;
 Onchocercidae; Dirofilaria.
 1 (bases 1 to 546)
 McCarter, J., Clifton, S., Chapell, B., Page, D., Martin, J.,
 Wylie, T., Dante, M., Maira, M., Hillier, L., Kucaba, T., Theising, B.,
 Bowers, Y., Gibbons, M., Ritter, B., Bennett, J., Franklin, C.,
 Tsagaris, V., Ronko, I., Kennedy, S., Maguire, L., Beck, C.,
 Underwood, K., Steptoe, M., Allen, M., Person, B., Swallow, T.,
 Harvey, N., Schurk, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
 McCann, R., Waterston, R. and Wilson, R.
 The Washington Univ. Nematode EST Project, 1999
 The Washington Univ. Nematode EST Project, 1999
 Contact: McCarter JP
 Unpublished (1999)
 The Washington Univ. Nematode EST Project, 1999
 The Washington Univ. Nematode EST Project, 1999
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.edu

TITLE
 JOURNAL
 COMMENT
 Possible reversed clone: similarity on wrong strand
 Seq primer: -40RP from Gibco
 High quality sequence stop: 418.
 Location/Qualifiers
 1..546
 /organism="Dirofilaria immitis"
 /mol_type="mRNA"
 /db_xref="taxon:6287"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Dirofilaria immitis adult PAMP1 v1"
 /note="Vector: PAMP1 (Gibco), Site 1: NotI, Site 2: SalI;
 The library was constructed by Claire Murphy and Dr. James McCarter
 at Washington University, St. Louis. Adult nematodes were harvested
 from infected dogs by Dr. Prema Arasu of North Carolina State
 University, Raleigh, NC (Prema.Arasu@ncsu.edu) Total RNA was
 isolated by Merry McAlaird of Divergence, Inc., St. Louis, MO. DNA
 Sequencing by: Washington University Genome Sequencing Center

FEATURES
 source
 1..546
 /organism="Dirofilaria immitis"
 /mol_type="mRNA"
 /db_xref="taxon:6287"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="Dirofilaria immitis adult PAMP1 v1"
 /note="Vector: PAMP1 (Gibco), Site 1: NotI, Site 2: SalI;
 The library was constructed by Claire Murphy and Dr. James McCarter
 at Washington University, St. Louis. The cDNA was
 made by using Dynabead Oligo-dT priming (Dyna1). PCR based
 synthesis kit from Clontech. Directionally cloned into the
 UDG sites of PAMP1. Adult nematodes were harvested from
 infected dogs by Dr. Prema Arasu of North Carolina State
 University, Raleigh, NC (Prema.Arasu@ncsu.edu). Total RNA

ORIGIN

was isolated by Merry McIaird of Divergence, Inc., St. Louis, MO."

Query Match 2.6%; Score 123.4; DB 13; Length 546;
Best Local Similarity 58.6%; Pred. No. 4.1e-09;
Matches 214; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

QY 953 AAAGTATCGAATTAATCACTCAATTTCCAGTAATAGCAGAAATTTTACAA 1012
DB 546 AAAGATATGAATGATCAATTAATTAAGCAAGAAATAGTAAGATTTTACT 487
QY 1013 TATATCCAAATTTTACTCTGTGCTCAAGTTTGTCTAAAGAAATTCAGAAATCTT 1072
DB 486 AACACCCCTGTGCTATTTTTCAGGCTTCTAGCTTTCAGATGAAAGTCGCAAAAAATG 427
QY 1073 CCTTGTATATAGTACTGCTGCTGATTAATAAGAACTTGATTCATGATGAAACA 1132
DB 426 CCTTATCAATGCTTCTGCTGATGCAAACTATACATGCTCAAGTTGATGCAACA 367
QY 1133 ATATGATATGATGTTTAAATTAATATATCAATCAAGATATATATGATGATGCA 1192
DB 366 CTATACCAAGAAATATCAAGTTTATAGTATGATTTATAGGAGTGCATTTTGT 307
QY 1193 GCTGATTAAGACATTAATTCATGATGATGATGATGATGATGATGATGATGAT 1252
DB 306 GCGACATTAAGACATTTTCTGATGATGATGATGATGATGATGATGATGATGAT 247
QY 1253 AATATGCTGTTGCTACTGTTATTAATTAAGGATGATGATGATGATGATGAT 1312
DB 246 TACAAATCCACGACGATGTTAGGCGATTAATTAATGACATGACATCTTAATTC 187
QY 1313 GCAAA 1317
DB 186 AGAAA 182

RESULT 22
CG758143/c 1297 bp DNA linear GSS 24-OCT-2003
LOCUS P053-3-B11.zb Ppa EcORI BAC Library Pristionchus pacificus genomic.
DEFINITION genomic survey sequence.

ACCESSION CG758143.1 GI:37987385
VERSION CG758143.1
KEYWORDS GSS.
SOURCE Pristionchus pacificus
ORGANISM Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Nemodiplogasteridae; Pristionchus.

REFERENCE 1 (bases 1 to 1297)
Strihvasan,J., Sinz,W., Jesse,T., Wiggers-Perebolle,L., Jansen,K.,
Bunjer,J., van der Meulen,M. and Sommer,R.J.
An integrated physical and genetic map of the nematode Pristionchus
pacificus

TITLE
JOURNAL Mol. Genet. Genomics 269 (5), 715-722 (2003)
MEDLINE 22835951
PUBMED 12884007

COMMENT
Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@uebingen.mpg.de

FEATURES
source
Class: BAC ends.
Location/Qualifiers
1..1297
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Ppa EcORI BAC Library"
/note="The library was generated by a partial digest of

ORIGIN

the genomic DNA with EcoRI and cloning into the BAC vector."

Query Match 2.6%; Score 123; DB 29; Length 1297;
Best Local Similarity 40.9%; Pred. No. 3.5e-09;
Matches 441; Conservative 0; Mismatches 636; Indels 0; Gaps 0;

QY 1276 AACTAAAGCATGAATGAATTAACACATATATATAGCAAAATCATTCATGATCT 1335
DB 1290 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1231
QY 1336 TCATACATTAATGCTCATCATGCTCTTGAGATCTTAATTAACATGACACAGACA 1395
DB 1230 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1171
QY 1396 TTCACGCAATATGCTTGTGACTAGAAATAGAAAGTGAATATATATCATTAAT 1455
DB 1170 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1111
QY 1456 AGATCAACCTTAAGCTTGTGAGAAACAGTACTGTAAACCACTGATATCATTAAC 1515
DB 1110 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1051
QY 1516 AAAAAAAAACTTAATGTAAGTACTACATTTGCAATCTATTTACATTTTACATGAGA 1575
DB 1050 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 991
QY 1576 TATATCACTGATTAAGCATATCAACATATATCTTATGCTATCATCTTTGATTAAT 1635
DB 990 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 931
QY 1636 TCAACCAATGATTAATTAAGCTCAAAATATGATCAACCTTCACTATATTA 1695
DB 930 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 871
QY 1696 AAAGCATCAATATCCATGTTTATACATCAATAGTATGATATTTATCAACAA 1755
DB 870 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 811
QY 1756 CTATAGTATCTTAATATTTCTTACTATATATTAACAAATCTTCTTCTATCTC 1815
DB 810 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 751
QY 1816 AATTAACAATGTTTATGATTAATTAATGATTAATTAATTTTATTA 1875
DB 750 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 691
QY 1876 ATATATGTTTGCACCTTTTCACTACATCATGATATGATGATATTTTACATA 1935
DB 690 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 631
QY 1936 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1995
DB 630 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 571
QY 1996 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2055
DB 570 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 511
QY 2056 AATATGATTAATCTGTTTTTATGCTTATGCTTATTAATTAATTTGCTATTTCCCA 2115
DB 510 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 451
QY 2116 AAAGGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2175
DB 450 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 391
QY 2176 TTAACCAATTTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2235
DB 390 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 331
QY 2236 CTATAGTATCTTCACTACATGATTAATTAATTAATTAATTAATTAATTAATTT 2295

| | | | | |
|---------------------------|--|---------------------------------------------------------------------|-----------------------------------------------------------------------|----------------------------|
| Dd | | 330 | AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA | 271 |
| Oy | | 2236 | ATGCAATTTCAGAAATTTCCACTTGCACACAAAACCTCAAAGTCTATTATATACADA | 2352 |
| Dd | | 270 | AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA | 214 |
| RESULT 23 | | | | |
| LOCUS | | CG747324/c | | |
| DEFINITION | | CG747324 | 1211 bp | DNA linear GSS 24-OCT-2001 |
| ACCSSION | | P040-4-D10.za Ppa EcORI BAC Library Pristionchus pacificus genomic | | |
| VERSION | | CG747324 | | |
| KEYWORDS | | CG747324.1 GI:37968250 | | |
| SOURCE | | GSS. | | |
| ORGANISM | | Pristionchus pacificus | | |
| REFERENCE | | Pristionchus pacificus | | |
| AUTHORS | | Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida; | | |
| | | Neodiplogasteridae; Pristionchus. | | |
| | | 1 (bases 1 to 1211) | | |
| | | Srinivasan,J., Stiz,M., Jesse,T., Wiggers-Perebolte,I., Jansen,K., | | |
| | | Bunjer,J., van der Meulen,M. and Sommer,R.J. | | |
| | | An integrated physical and genetic map of the nematode Pristionchus | | |
| | | pacificus | | |
| JOURNAL | | Mol. Genet. Genomics 269 (5), 715-722 (2003) | | |
| MEDLINE | | 22835951 | | |
| PUBMED | | 12884007 | | |
| COMMENT | | Contact: Sommer RJ | | |
| | | Evolutionary Biology | | |
| | | Max-Planck-Institute for Developmental Biology | | |
| | | Spemannstr. 37-39, Tuebingen D-72076, Germany | | |
| | | Tel: 00497071601371 | | |
| | | Fax: 00497071601498 | | |
| | | Email: ralf.sommer@uebingen.mpg.de | | |
| | | Class: BAC ends. | | |
| FEATURES | | | | |
| source | | Location/Qualifiers | | |
| | | 1..1211 | | |
| | | /organism="Pristionchus pacificus" | | |
| | | /mol_type="genomic DNA" | | |
| | | /strain="California" | | |
| | | /db_xref="taxon:54126" | | |
| | | /clone_lib="Ppa EcORI BAC library" | | |
| | | /note="The library was generated by a partial digest of | | |
| | | the genomic DNA with EcORI and cloning into the BAC | | |
| | | vector." | | |
| ORIGIN | | | | |
| | | | | |
| Query Match | | 2.6%; Score 122.8; DB 29; Length 1211; | | |
| Best Local Similarity | | 39.9%; Pred. NO. 3.8e-09; | | |
| Matches 427; Conservative | | 0; Mismatches 643; Indels 0; Gaps 0; | | |
| Oy | | 1496 | AAACGACGTGATATCTTAGCAGAAAAAAGCTTAATGTGAACAATTCATATCTATT | 1555S |
| Dd | | 1209 | AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA | 1150S |
| Oy | | 1556 | TACATTTATTACATGAGATATATCATGATTAAGCCATATCAACATATTTCTTAG | 1615S |
| Dd | | 1149 | AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA | 1090S |
| Oy | | 1616 | TCTATCATCCTTTGATAATTTCCACACATGATAATTAAGGTCAGAAATATATGAAATAC | 1675S |
| Dd | | 1089 | NAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA | 1030S |
| Oy | | 1676 | AAGGTCACCATATATTAAGAAGAGTCAATGCCATCGTTACATCCAATCACTAGTAT | 1735S |
| Dd | | 1029 | AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA | 970 |
| Oy | | 1736 | GTAACATTTTCATPAACAACCTTAAGTTAATCTTAATTTCCAGCTAAATTATATTAAC | 1795S |
| Dd | | 969 | AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA | 910 |
| Oy | | 1796 | AAATTCCTATCTCATCTCAATTAACCAATGGTATTCAGATTATTAATAGTCAATTA | 1855S |

[illegible]

Email: submissions@watscn.wustl.edu
 Insert Length: 18200 Std Error: 0.00
 Seq primer: Sp6 ATTGAGTACACATATAG
 Class: BAC ends
 High quality sequence start: 34
 High quality sequence stop: 261.

FEATURES

source

1.1272
 Location/Qualifiers
 /organism="Gallus gallus"
 /mol_type="genomic DNA"
 /strain="Red Jungle Fowl"
 /db_xref="taxon:9031"
 /clone="CH261-19L17"
 /sex="female"
 /cell_line="UCD001, inbred 256"
 /clone_1ib="CH261"
 /note="Vector: pRABAC2.1; Site_1: EcoRI; Site_2: EcoRI;
 CH261 Female Chicken library - for library and clone
 ordering information: http://www.choxi.org/bacpac"

ORIGIN

Query Match 2.6%; Score 121.8; DB 28; Length 1272;

Best Local Similarity 47.7%; Pred. No. 5.3e-09;

Matches 407; Conservative 0; Mismatches 444; Indels 3; Gaps 2;

1604 ATATTATCTAGCTATCATCTCTTGATTAATTTCAACA-CATGATTAATTAAGGTCAA 1662
 1236 ATTAATATATAGTTAATATATATATATATATTAATTAATTAATATATATATATAT 1177
 1663 AAATAGATCAACAGTCCACCATATATTAAGAGAGCAATACCCCTCGTTACA 1722
 1176 TAAATATATATATATATATATATATATATATATATATATATATATATATATAT 1117
 1723 TCCAACTAGTATGTAACATATATATATATATATATATATATATATATATATATAT 1782
 1116 AAAAAATATATATATATATATATATATATATATATATATATATATATATATATAT 1057
 1783 TATATATATTAACAATTCCTATCTCTCATCTCAATTAACAATTTGATCATATATA 1842
 1056 AATATATATATATATATATATATATATATATATATATATATATATATATATAT 997
 1843 ATATAGCAATTAACCTTTTACTTATTTATTAATTAATTAATTTGACACCTTTCACTA 1992
 996 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 937
 1903 CATCATGATATATATATATATATATATATATATATATATATATATATATATATAT 1962
 936 TATATATATATATATATATATATATATATATATATATATATATATATATATATAT 879
 1963 ATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2022
 878 ATATATATATATATATATATATATATATATATATATATATATATATATATATAT 819
 2023 AATTCATCATCTCTGATATTTCAACAATTAATTAATTAATTTGATTTTATGCC 2082
 818 TATATATATATATATATATATATATATATATATATATATATATATATATATATAT 759
 2083 CATTCGTTCTATATATATTTGCTATTTACCAAAAAGTTAAATACCATCTCTGACAC 2142
 758 TATTTATATATATATATATATATATATATATATATATATATATATATATATATAT 699
 2143 AATATAGAGAGATACCAATACAGTCAATTAATTTAAACAATTTCTATATTAACAACCT 2202
 698 AATATATATATATATATATATATATATATATATATATATATATATATATATATAT 639
 2203 ACTTAAACATAGAGACCAATAGTATATCATACCAATAGATATCCCTACACATGTA 2262
 638 TTAATATATATATATATATATATATATATATATATATATATATATATATATATAT 579
 2263 TCAATATATATATATATATATATATATATATATATATATATATATATATATATAT 2322
 578 AAAAAATATATATATATATATATATATATATATATATATATATATATATATATAT 519

QY 2323 CAACAAAACCTCAAGTCTTATATATTAACAATTCAGAAATTATATATCTTCAACCAACA 2382
 DB 518 TAAATATATATATATATATATATATATATATATATATATATATATATATATATAT 459
 QY 2383 TACTATCAATCTCTATATATATATATATATATATATATATATATATATATATATAT 2442
 DB 458 AATTTATATATATATATATATATATATATATATATATATATATATATATATATAT 399
 QY 2443 CATTAATATATATATATATATATATATATATATATATATATATATATATATAT 2456
 DB 398 TAAATATATATATATATATATATATATATATATATATATATATATATATATAT 385

RESULT 25

CG753854/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

1 (bases 1 to 1528)
 Srinivasan, J., Sins, W., Jesse, T., Wiggere-Perreolte, L., Jansen, K.,
 Buntjer, J., van der Meulen, M., and Sommer, R.J.
 An integrated physical and genetic map of the nematode *Pristionchus*
pacificus.
 Mol. Genet. Genomics 269 (5), 715-722 (2003)
 Contact: Sommer, R.J.
 Evolutionary Biology
 Max-Planck-Institute for Developmental Biology
 Spemannstr. 37-39, Tuebingen D-72076, Germany
 Tel: 00497071601371
 Fax: 00497071601498
 Email: ralf.sommer@tuebingen.mpg.de
 Class: BAC ends.
 Location/Qualifiers
 1.1528
 /organism="Pristionchus pacificus"
 /mol_type="genomic DNA"
 /strain="California"
 /db_xref="taxon:54126"
 /clone_1ib="Ppa EcoRI BAC library"
 /note="The library was generated by a partial digest of
 the genomic DNA with EcoRI and cloning into the BAC
 vector."

FEATURES

source

QY 1276 AACTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1335
 DB 1485 AANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1426
 QY 1336 TCATACATTAATTTGTCATCATCTCTTGAGATCTAATATTACATGATACAGACACA 1395
 DB 1425 AANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1366
 QY 1396 TTCACCAATATAGGCTTTTGCACTAGAAATAGGAGGAGGAGGAGGAGGAGGAGG 1455
 DB 1365 AANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1306
 QY 1456 AGATCAACAACCTTAAGCTTTGTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1515
 DB 1305 AAAAAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1246

[illegible]

| | |
|--------------------------|----------------------------------------|
| RESULT 26 | |
| CG744812/c | |
| LOCUS | 1353 bp |
| DEFINITION | DNA linear |
| CG744812 | SSS 24-OCT-2003 |
| P037-3-B03_Ya Ppa EcORI | BAC library |
| Genomic survey sequence. | <i>Pristionchus pacificus</i> genomic. |
| CG744812 | |
| ACCESSION | |
| VERSION | |
| CG744812.1 | GI:37965680 |
| KEYWORDS | |
| SSS. | |

| ORGANISM | SOURCE |
|----------------------------------------------------------------------------|----------------------------------------|
| Pristionchus pacificus | Pristionchus pacificus |
| Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida; | |
| Neodiplogasteridae; Pristionchus. | |
| 1 (baes 1 to 1353) | |
| Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolle,L., Jansen,K., | |
| Buntjer,J., Van der Meulen,M. and Sommer,R.J. | |
| An integrated physical and genetic map of the nematode <i>Pristionchus</i> | |
| <i>pacificus</i> | |
| Mol. Genet. Genomics 269 (5) , 715-722 (2003) | |
| 22835951 | |
| 12884007 | |
| Contact: Sommer R.J | |
| Evolutionary Biology | |
| Max-Planck-Institute for Developmental Biology | |
| Spemannstr. 37-39, Tuebingen D-72076, Germany | |
| Tel.: 00497071601371 | |
| Fax: 00497071601498 | |
| Email: ralf.sommer@uebingen.mpg.de | |
| Class: BAC ends. | |
| Location/Qualifiers | |
| 1..1353 | |
| /organism="Pristionchus pacificus" | |
| /mol_type="genomic DNA" | |
| /strain="California" | |
| /db_xref="taxon:54126" | |
| /clone_lib="Ppa Scori BAC Library" | |
| /note="The library was generated by a partial digest of | |
| the genomic DNA with EcoRI and cloning into the BAC | |
| vector." | |
| ORIGIN | |
| Query Match | 2.5%; Score 120.6; DB 29; Length 1353; |
| Best Local Similarity | 39.6%; Pred. No. 7.9e-09; |
| Matches 422; Conservative 0; Mismatches 643; Indels 1; Gaps 1; | |
| QY 1420 AGAATATGGAAGAGTGAATATTAATACATTATATGATCAACAACCTAAAGCTTGTGA 1479 | |
| DB 1270 AAAAAANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1211 | |
| QY 1480 AGGAACACGTCGTAAACACGATATATCGATTTAGCAAAAAAACTTAATGTAGAACTAC 1539 | |
| DB 1210 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1151 | |
| QY 1540 AATTGCATATCTATTTACATTTATTACATGCAATATATCACTAGATAAAGCATATC 1599 | |
| DB 1150 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANN 1091 | |
| QY 1600 AAAGATTTATCTAGTCTATCATCCCTTGATATT-CAACGAATGATTAATTAACG 1658 | |
| DB 1090 AAAAAANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1031 | |
| QY 1659 TCAAAAATATGAAATACAGCTTCACCAATATATAAAAGAGTCAATACCATGGT 1718 | |
| DB 1030 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAN 971 | |
| QY 1719 AACGTCACACTAGATATGTAATATATACATACAACTTAAGTATACCTTAATTTCC 1778 | |
| DB 970 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 911 | |
| QY 1779 TAGCTAATTTATATAACAATTCCTATCTCTCATCTCAATTAACCAATGTGTACAGAT 1838 | |
| DB 910 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 851 | |
| QY 1839 ATAAATAGTCAATTAACCTTTACTATTATTAATAATATATGTTGCACACTTTTC 1898 | |
| DB 850 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANA 791 | |
| QY 1899 ACTACATCATGTATATATGACATTAATTTTACATAAATATACATATTAAGAACCAAT 1958 | |
| DB 790 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 731 | |
| QY 1959 ACATATATTAATTTACTAAACACCTATTACTACGTAATATATACATGATGTCCTGACA 2018 | |

QY 2388 ATCAATTCATATACATCATCTTTTAAAGAGAGCTGCACATTCATACATA 2447
 DB 1188 TATATATATATATATATATATATATATATATATATATATATATATATAT 1247
 QY 2448 ATATATATATATATATATATATATATATATATATATATATATATATAT 2477
 DB 1248 ATATATATATATATATATATATATATATATATATATATATATATATAT 1277

RESULT 28

CG745119/c 1433 bp DNA linear GSS 24-OCT-2003
 LOCUS P037-4-G03.ya Ppa EcORI BAC Library Pristionchus pacificus genomic,
 DEFINITION genomic survey sequence.
 ACCESSION CG745119
 VERSION CG745119.1 GI:37966045
 KEYWORDS GSS
 SOURCE Pristionchus pacificus
 ORGANISM Pristionchus pacificus
 Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
 Neodiplogasteridae; Pristionchus.

REFERENCE

AUTHORS Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
 Buntjer,J., van der Meulen,M. and Sommer,R.J.
 An integrated physical and genetic map of the nematode Pristionchus
 pacificus
 Mol. Genet. Genomics 269 (5), 715-722 (2003)

TITLE

JOURNAL Evolutionary Biology
 MEDLINE Max-Planck-Institute for Developmental Biology
 PUBMED Spermastoc. 37-39, Tuebingen D-72076, Germany
 COMMENT Tel: 00497071601371
 Fax: 00497071601498
 Email: ralf.sommer@uebingen.mpg.de
 Class: BAC ends.

FEATURES

source Location/Qualifiers
 1..1433
 /organism="Pristionchus pacificus"
 /mol_type="genomic DNA"
 /strain="California"
 /db_xref="taxon:54126"
 /clone_lib="Ppa EcORI BAC Library"
 /note="The library was generated by a partial digest of
 the genomic DNA with EcORI and cloning into the BAC
 vector."

ORIGIN

Query Match 2.5%; Score 120.2; DB 29; Length 1433;
 Best Local Similarity 41.4%; Pred. No. 8.9e-09;
 Matches 457; Conservative 0; Mismatches 646; Indels 1; Gaps 1;
 QY 1275 TAACATGAGCATGATGAAATTAAGACATATATATGCAAAAAATCATTCATGATC 1334
 DB 1423 TAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1364
 QY 1335 TTCAATCATTAATGCTCATGCTCTTGAGATCTAATATTAACATGTACACAGAAC 1394
 DB 1363 TAAATAAANTANAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1304
 QY 1395 ATTCAAGCAATATGCTTTTGACTAGAAATAGAAAGGTAAGATATTAATGATTA 1454
 DB 1303 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1244
 QY 1455 TAGATCAACACTTAAGCTGTGAGGACCACTGATCTGTAACCACTGATTCATTAG 1514
 DB 1243 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1184
 QY 1515 CAAAAAACTTAATGATGCAATCTCAATTGG-CATATCTATTTCATTAATTAATGATG 1573
 DB 1183 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1124

QY 1574 AATATATCACTGATTAAGCAGATATCAACATATTAATCTTAGCTATCATCTCTTGATA 1633
 DB 1123 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1064
 QY 1634 TTCAACACAGATGATTAATTAACGGTCAAAAAATATGATATACAGGTTCCACCATATAT 1693
 DB 1063 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1004
 QY 1694 AAAAAAGCTCAATATCCCATGCTTATACATCCATGATATGATATATATATATATAT 1753
 DB 1003 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 944
 QY 1754 AACTTAGATATATCTTAATATTTCTAGCTAATTAATTAACAAATTCTATCTCTATC 1813
 DB 943 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 884
 QY 1814 TCAATTAACCAATGTTTATGATTTAATTAAGTCAATTAACCTTTATCTTATATTA 1873
 DB 883 NNNNAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 824
 QY 1874 AATATATATGTTTGACACCTTTTCTACATCATGATATATATGACATATTTTACAT 1933
 DB 823 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 764
 QY 1934 AAAAAATCAATTAATTAAGACCAATATATATTAATTAATTAATTAATTAATTAAT 1993
 DB 763 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 704
 QY 1994 GATATATTAATGATGCTCTGACATTAATTAATTTATCATCTCTGACATATTTTCAACA 2053
 DB 703 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 644
 QY 2054 TAAATATGATTAATGATGTTTATGATGCTTATTAATTAATTAATTAATTAATTA 2113
 DB 643 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 584
 QY 2114 AAAAAAGTTAAATTAATGATGCTCTGACATTAATTAATTAATTAATTAATTAAT 2173
 DB 583 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 524
 QY 2174 AATTAACCAATTTCTAATTAATTAACCACTCTTAACATAGAAACCAATATATATCAT 2233
 DB 523 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 464
 QY 2234 ACTTAATGATATATCCCTGACGATGATCAATTAATTAATTAATTAATTAATTAAT 2293
 DB 463 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 404
 QY 2294 TTAATCAATTAACCAATTTCCACTTCAACCAACCAATCAAGTTCTATTAATCAAT 2353
 DB 403 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 344
 QY 2354 TCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2377
 DB 343 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 320

RESULT 29

CG750135 1592 bp DNA linear GSS 24-OCT-2003
 LOCUS P044-3-D09.za Ppa EcORI BAC Library Pristionchus pacificus genomic,
 DEFINITION genomic survey sequence.
 ACCESSION CG750135
 VERSION CG750135.1 GI:37971278
 KEYWORDS GSS
 SOURCE Pristionchus pacificus
 ORGANISM Pristionchus pacificus
 Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
 Neodiplogasteridae; Pristionchus.

REFERENCE

AUTHORS Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
 Buntjer,J., van der Meulen,M. and Sommer,R.J.
 An integrated physical and genetic map of the nematode Pristionchus

Db 338 AAAAAAAAAAANNANNTA--ACATACCATCTACCTCTCTTAACATA 281

Qy 2354 TGAAGATATATATCTTACCAACATAGTATCAATCTTATATACATATCC 2413

Db 280 CCAATTTACCCATATNCCACCTATNTTTCTCTCTCAATCAAAAANTTAACA 221

Qy 2414 TTTTAAAGAGTCTTACATTCATATCAATTAATTAATGATGCAATGTTATCA 2473

Db 220 TCAACCATCTCTNTTNCACATCTTATNATACCACTNCAACAACTNCAATN 161

Qy 2474 GAAACATTACATCAAGAAATTT 2499

Db 160 TNAACCTCAAAATTTTCAATAT 135

RESULT 33

LOCUS BX436282/c 1124 bp mRNA linear EST 15-MAY-2003

DEFINITION BX436282 Homo sapiens THYMUS Homo sapiens cDNA clone CSOCAP001YC01

ACCESSION BX436282

VERSION 5-PRIME, mRNA sequence.

KEYWORDS EST.

SOURCE BX436282.1 GI:30787521

ORGANISM Homo sapiens (human)

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

AUTHORS Li W.B., Gruber C., Jessee J. and Polayes D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)

COMMENT Contact: Genoscope

Genoscope - Centre National de Sequencage

Bp 191 91006 EVRY cedex - France

Email: segre@genoscope.cns.fr; Web: www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of Invitrogen. Contact: Feng Liang Email: fliang@life.techn.com URL: http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID: CSOCAP001A010P1.

FEATURES

source

1..1124

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CSOCAP001YC01"

/csize_type="THYMUS"

/clone_lib="Homo sapiens THYMUS"

/note="Vector: pCMVSPORT 6; 1st strand cDNA was enriched with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

Query Match 2.5%; Score 117.8; DB 13; Length 1124;

Best Local Similarity 27.1%; Pred. No. 2.2e-08;

Matches 215; Conservative 255; Mismatches 308; Indels 16; Gaps 2;

Qy 1564 ATTACATGGAATATATCAATGAATTAAGCATATCAACATTAATCTTATCTATCAT 1623

Db 1121 AAAAAAAAAAAAAAAAAAHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH 1062

Qy 1624 CTTTGATATTTTCAACATGATGATTAATTAACGGTCAAAATATAGATGATCAACGTCA 1663

Db 1061 YNYYHHHAAAYYYAAYYYHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH 1002

Qy 1664 CCACATATATTAAGAGCATGCAATCCATGTTAATCAATCAACTAGATATGTTACATA 1743

Db 1001 YYYYYYAAAAAAAAAAAAAAAAHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH 942

Qy 1744 TTAGCATCAAACTTAATGATTAATTAATTTCTCTGATCTATATATTAACAAATTCCT 1803

Db 941 YYYYYYYYYYYYYYYYYYHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH 882

Qy 1804 ATCTCTATCTCAATTACCAATTTGTTATCAAGTTATTAATTAAGTC-----A 1850

Db 881 YYYYYHHHAAAAAAAAAYYYHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH 822

Qy 1851 AATTAACCTTTATATTTATTAATAATATATGTTGACACTTTTCACTACATCATGT 1910

Db 821 AAAAAAYYYYYYYYYYAAAAAAAAAAAAAAAAAYYYAAAYYYAAAYYYAAAYYYAA 762

Qy 1911 ATATATGAACATATTTTTCATATAATACTATATAAAGAACCAATCATATATTAA 1970

Db 761 MHHHAAAAAAAAAAHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH 702

Qy 1971 TTAATAAACCCTATTACTATGCTAATTAATTAATGATGCTGACATATATATCTA 2030

Db 701 YYYYYYYYYYYYYYYYYYHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH 642

Qy 2031 TCATCTTGACATATTTCAACATTAATAATGATAAGTGTTTTATGCCCATTAAGT 2090

Db 641 YYYYYYYYYHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH 582

Qy 2091 CTTAATTAATTTGCTATATACCAAAAGTTAAATTAATCAACATCTGCAATATGA 2150

Db 581 YYYYYHHHAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAHHHHHHHHHHHHHH 522

Qy 2151 GAAGATACATACAGCTCAATTAATTAACAAATTTCTAATTAACCACTCTCTAA 2210

Db 521 AAAAAAAAAAAAAA---HAAAAAAAAAAAYAAHHHHHHHHHHHHHHHHHHHHHH 465

Qy 2211 CATAGGACCATATGATATCATCTAAGTAAGATATCCCTACAGCATGATCAATTA 2270

Db 464 AAAAAAAAAAAAYHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH 405

Qy 2271 TAAAGAAAAATCGTATGATATTTATTAATTAATTAATTTCAATCTCAACAAATA 2330

Db 404 YHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH 345

Qy 2331 CTCAAGTCTTAT 2344

Db 344 YYYAAHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH 331

RESULT 34

LOCUS CC231597/c 1426 bp DNA linear GSS 12-MAY-2003

DEFINITION CH261-36A19 RM1.1 CH261 Gallus gallus genomic clone CH261-36A19,

genomic survey sequence.

ACCESSION CC231597

VERSION CC231597.1 GI:30558260

KEYWORDS GSS.

SOURCE Gallus gallus (chicken)

ORGANISM Gallus gallus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.

AUTHORS Krenitzki C., Higginbotham J., Wylie K., Carter J., McPherson J., Warren W., Graves T., Mardis E. and Wilson R.

TITLE Gallus gallus BAC End Reads

JOURNAL Unpublished (2003)

COMMENT Contact: Richard K. Wilson

Genome Sequencing Center

Washington University School of Medicine

Email: submissions@wustl.wustl.edu

Insert Length: 18200 Std Error: 0.00

Seq primer: RM1 TAGCATCATATAGGGAGA

Class: BAC ends

High quality sequence start: 276

High quality sequence stop: 339.

FEATURES

source

1..1426

/organism="Gallus gallus"

/mol_type="genomic DNA"

/strain="Red Jungle Fowl"
 /db_xref="taxon:9031"
 /clone="CH261-36A19"
 /sex="female"
 /cell_line="UCD001, inbred 256"
 /note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;
 CH261 Female Chicken library - for library and clone
 ordering information: <http://www.chori.org/bacpac>"

ORIGIN

Query Match Best Local Similarity 41.3%; Score 117.8; DB 28; Length 1426;
 Matches 416; Conservative 0; Mismatches 591; Indels 0; Gaps 0;

1378 AACATGTACACGAAACATTCACGCAATTCGCTTTGGACATGAAATAGAAAAGTAG 1437
 1388 AAGAAAGAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAG 1329
 1438 AATATATATCTTATATGATGACACCTAAAGCTTGTGAAGAACAGTACTGTAA 1497
 1328 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1269
 1498 ACCAGTATATCATTCAGCAAACTTATGTAGACCTACCAATTTGCATATCTATTA 1557
 1268 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1209
 1558 CAATTTATATCATGAGATATATCTAGATATAGCCATATCAACATATATCTTACTG 1617
 1208 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1149
 1618 TATCATCTTGTATATTTCAACACATGATTAATTAAGGTCAAAATATAGATTA 1677
 1148 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1089
 1678 CGTTCACACATATATTAAGACAGTCAATATCCATGTTACATCCAGTATAGT 1737
 1088 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1029
 1738 AACATATTCATACAACTTAAAGTATATCTTAATTTCCAGTATATATTAACA 1797
 1028 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 969
 1798 ATTCATCTCTATCTCATTCATTAACCAATGTTATAGATTAATTAAGTCAATTA 1857
 968 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 909
 1858 CTTTATCTTATTAATTAATATATGTTTGCACACTTTTCACTACATCATATATAT 1917
 908 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 849
 1918 GAACATATTTTACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1977
 848 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 789
 1978 AACACCTATTTATCTAGTATATTAATCATAGATGCTGACATATTAATTTCTATCT 2037
 788 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 729
 2038 TGACATATTTCAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2097
 728 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 669
 2098 AATATGCTATTAACCAAAAGTTAAATACACATTCCTGCAATATGAGAGATA 2157
 668 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 609
 2158 CAATACAGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2217
 608 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 549
 2218 ACCAATATATATCTATCTATAGATATCTTACATGATCTCAATTAATTAAGAA 2277

Db 548 NAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 489
 Qy 2278 AATTCGTATGTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2337
 Db 488 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 429
 Qy 2238 TTCTATATTAACAAATTCAGAAATTAATTAATTAATTAATTAATTAATTAATTA 2384
 Db 428 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 382

RESULT 35

CC293489 1596 bp DNA linear GSS 13-MAY-2003
 LOCUS CH261-62N1.RM1.2 CH261 Gallus gallus genomic clone CH261-62N1,
 DEFINITION genomic survey sequence.
 ACCESSION CC293489
 VERSION CC293489.1 GI:30664930
 KEYWORDS GSS.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 1 (bases 1 to 1596)
 Krenitzki, C., Higginbotham, J., Wylie, K., Carter, J., McPherson, J.,
 Warren, W., Graves, T., Mardis, E. and Wilson, R.
 Gallus gallus BAC End Reads
 Unpublished (2003)
 CONTACT: Richard K. Wilson
 GENOME SEQUENCING CENTER
 WASHINGTON UNIVERSITY SCHOOL OF MEDICINE
 EMAIL: submissions@wustl.edu
 INSERT LENGTH: 182000 Std Error: 0.00
 SEQ PRIMER: RM1 TAGACTCCTATAGGAGA
 CLASSE: BAC ends
 HIGH QUALITY SEQUENCE START: 1001
 HIGH QUALITY SEQUENCE STOP: 1061.
 LOCATION/Qualifiers
 1. 1596
 /organism="Gallus gallus"
 /mol_type="genomic DNA"
 /strain="Red Jungle Fowl"
 /db_xref="taxon:9031"
 /clone="CH261-62N1"
 /sex="female"
 /cell_line="UCD001, inbred 256"
 /note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;
 CH261 Female Chicken library - for library and clone
 ordering information: <http://www.chori.org/bacpac>"

FEATURES

Source

1. 1596
 /organism="Gallus gallus"
 /mol_type="genomic DNA"
 /strain="Red Jungle Fowl"
 /db_xref="taxon:9031"
 /clone="CH261-62N1"
 /sex="female"
 /cell_line="UCD001, inbred 256"
 /note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;
 CH261 Female Chicken library - for library and clone
 ordering information: <http://www.chori.org/bacpac>"

ORIGIN

Query Match Best Local Similarity 41.4%; Score 117.4; DB 28; Length 1596;
 Matches 490; Conservative 0; Mismatches 693; Indels 0; Gaps 0;

1274 ATAACTAAGGCGATGAATTAATTAACATATATATGCAAAATATCATTAATAGAT 1333
 73 AAAAAAAAAAAAAAAAAAGGCGCAAAAGGNNNNCCAAAAATTAAGAAAGAAANNNNN 1332
 1334 CTTCAATCAATTAATGCTCATCATCTCTTGAGATCTAATTAATCAATGTCACACAGA 1393
 1333 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCCAAATTAAGAAAGAAANNNNN 192
 1394 CATTCACGAATATGCTTTTGGACTAGAAATAGAGAAAGTAGAAATTAATATACATTA 1453
 193 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 252
 1454 ATAGATCACACCTAAAGCTTGTGAAGAACAGTACTGTAAACCACTGATATCATTA 1513
 253 AAAAAAAAAAAAAAAAAANCAACAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 312

QY 1514 GCAAAAACTTAATAGAACACCAATTGTCATATCTATTACATTTATTACATGAG 1573
 DB 313 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 372
 QY 1574 AATATATCATCTAGATTAAGCCATATCAACATATTTATCTTATGCTATCATCTTTGATTA 1633
 DB 373 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 432
 QY 1634 TTTCACACATGATTAATTAACGGTCAAAATATAGATTAACAGTTTCCACATATAT 1693
 DB 433 AGAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 492
 QY 1694 AAAAAAGCATCAATTAACCATGTTTAACTCACTAAGTATGTAACATTTATACATA 1753
 DB 493 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 552
 QY 1754 AACTTAAGTATTAATTAATTTCTAGTATTTATTAACAAATTTCTTCTCTATC 1813
 DB 553 AACAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 612
 QY 1814 TCATTAACCAATTTGTTATCAGATTATTAATTAAGTCAATTAACCTTTTATTTATTA 1873
 DB 613 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 672
 QY 1874 AATTAATATGTTGCACTTTTTCATCATCATGATATATATGACATTAATTTTACAT 1933
 DB 673 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 732
 QY 1934 AAAAAATCAATTAATAAGAACATATATTAATTAATTAATTAATTAATTAATTAATTA 1993
 DB 733 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 792
 QY 1994 GTAATATTAATGATGTCCTGACATATTAATTTATCATCTTTGATTAATTTCAACA 2053
 DB 793 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 852
 QY 2054 TAAATATATGATTAATGTTTATGTCCTATGTTCTATTAATTAATTTGCTAATTA 2113
 DB 853 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 912
 QY 2114 AAAAAAGTTAAATACATCTCTGCAACATATGTAAGATATACATACACATCAATA 2173
 DB 913 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 972
 QY 2174 AATTAACCAATTTCTATATATACCACTCTTAACATTAAGAACATATGATATCAT 2233
 DB 973 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 1032
 QY 2234 ACCTAAGTATGATTCCTGACGATGATATCAATTAATTAAGAAAAATCGTTATGTA 2293
 DB 1033 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 1092
 QY 2294 TATATCAATTAACCAATTTCTCACTCAACACAAATCTCAAGTTCTTATTAACAAT 2353
 DB 1093 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1152
 QY 2354 TCAAGATATATATATATCTACCAACATAGTATTAATTTCTATATATATATATAT 2413
 DB 1153 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 1212
 QY 2414 TTTTAAAAAGATGATTCACATTCATATCAATTAATTAATTAATTAATTAATTA 2456
 DB 1213 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 1255

RESULT 36

CNS033GQ

Tetradon nigroviridis genome survey sequence PUC-Ort end of clone

DEFINITION

20824 of library G from Tetradon nigroviridis, genomic survey

ACCESSION AL226115.1 GI:7885026

KEYWORDS

GSS: genome survey sequence.
Tetradon nigroviridis

SOURCE

Tetradon nigroviridis

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes; Tetraodontidae; Tetraodontidae; Tetraodon.

REFERENCE

1. Reest Crollius, H., Jallion O., Dasilva, C., Bouneau, L., Fisher, C., Bernot, A., Fzames, C., Wincker, P., Brothier, P., Quetier, F., Saurin, W., and Weissenbach, J.

TITLE

Estimate of human gene number provided by genome-wide analysis using Tetradon nigroviridis DNA sequence

JOURNAL

Nat. Genet. 25 (2), 235-238 (2000)

MEDLINE

20296633

PUBMED

10835645

REFERENCE

2. Reest Crollius, H., Jallion O., Dasilva, C., Ozouf-Coataz, C., Fzames, C., Fischer, C., Bouneau, L., Billaud, A., Quetier, F., Saurin, W., Bernot, A., and Weissenbach, J.

TITLE

Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetradon nigroviridis

JOURNAL

Genome Res. 10 (7), 939-949 (2000)

MEDLINE

20358837

PUBMED

10899143

REFERENCE

3. (bases 1 to 1135)
Genoscope.
Direct Submission
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage: BP 191 91006 EVRY cedex - FRANCE (E-mail: seqref@genoscope.cns.fr
- Web: www.genoscope.cns.fr)
This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetradon nigroviridis genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetradon.

FEATURES

Location/Qualifiers

source

1..1135
/organism="Tetradon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="20824"
/clone_1ib="G"
/note="Genoscope sequence ID : CNS033GQ125P1-end : PUC-Ort1"

ORIGIN

Query Match 2.5%; Score 117; DB 29; Length 1135;
Best Local Similarity 45.5%; Pred. No. 2.9e-08;
Matches 429; Conservative 15; Mismatches 486; Indels 10; Gaps 2;

QY 1507 ATCAATTAAGCAAAAACTTAATGATAGACTACCAATTTGCAATTCATTTTACATTTATT 1566
 DB 199 ATATATTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 258
 QY 1567 ACATGGAATATATCTAGATTAAGCATATCAACATATATATCTTATGATCACTCT 1626
 DB 259 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 318
 QY 1627 TTGATATTTCAACACATGATTAATAACGGTCAAAATATAGAAATACACAGTTCCACA 1686
 DB 319 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 378
 QY 1687 CATATTAATAAGCATCAATTAACCATGTTTAACTCACTCACTAAGTATGTAACATATTA 1746
 DB 379 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 438
 QY 1747 CATTAACCACTTAAGTATTAATTAATTTCTAGTAAATTAATTAACCAATTTCTATC 1806
 DB 439 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 498
 QY 1807 TCTATCTCAATTAACCAATTTGTTATGATATTAATTAATTAAGTCAATTAATTTTACT 1866
 DB 499 AATTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 558

| | | | |
|----|------|--------------------------------------------------------------|------|
| QY | 1867 | ATTCTAAATTAATATGTTGGACCTTTTCACAGCATCAGTATATATGACATATAT | 1922 |
| Db | 559 | ATTAAATTAATAAAAAAAAAAAAAAAAAATTAATTA-AAAAAAAAAAAAAAAAAA | 616 |
| QY | 1927 | TTTACATAAATTAAGTAAATTAAGAGACCATATATATTAATTTACTTAAACACTAT | 1986 |
| Db | 617 | TTTAAAAAATAAAAAAAAAAAAAAAAAATTAATAATTTAAAAAATAATTAATTTA | 676 |
| QY | 1987 | TTACTACGTAATATTACATAGATGTCCTGACATAATTAATTCATCTTGACATATT | 2046 |
| Db | 677 | AAAAAAAAAAAAATTAATTAATAAAAAAAAAAAAAAAAAATTAATAAAAAAAAAAAAA | 736 |
| QY | 2047 | TCAAACATAAATATGATTAACCTGTTTTTTATGCCATATAGTCTTAATTAATGCTA | 2106 |
| Db | 737 | AATTCAATTAATAAAAAAAAAAAAAAAAAATTCATAACAATTAATAAAAAAAAAATTAAT | 796 |
| QY | 2107 | ATTACCCAAAAGTTAAATATACATCTCTGCAACAATATGAGAGATACCAATACAG | 2166 |
| Db | 797 | AAMAYATTAATTAATAAAAAAAAAAACAACAACAACATTAATAAAAAAAAAAAAAAA | 856 |
| QY | 2167 | TCAATTAATTAACAAATTTCTAATATTAACCACTACTCTTAACATAGAAACATAGT | 2226 |
| Db | 857 | AAAAAAAAAAAAATAAAAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTA | 916 |
| QY | 2227 | ATATCATACCTAGTAAAGTATCCCTAAGACATGATCAATTAATTAAGAAAAATCGTA | 2286 |
| Db | 917 | AAA-----AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA | 968 |
| QY | 2287 | TGTATATTATTAACAATTATCAAAATTTCCAACTTCACACAAACTCCAAATGCTTATAT | 2346 |
| Db | 969 | AAAAATTAATAAAAAAAAAAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTA | 1026 |
| QY | 2347 | AACCAATTCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA | 2406 |
| Db | 1029 | AAAAAAAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA | 1088 |
| QY | 2407 | CATATCCTTTTAAAGATGCTTCAACAATTCACACATTAATTAATTAATTAATTAATTA | 2466 |
| Db | 1089 | AAAAAAAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA | 1130 |

```

/molecule="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_id="Pa EcoRI BAC Library"
/notes="The library was generated by a partial digest of
the genomic DNA with EcoRI and cloning into the BAC
vector."

```

| | Query Match | 2.5#: | Score 117: | DB 29: | Length 1269: | | | |
|----|-----------------------|----------|--------------------|----------|-----------------------------------|---------------------|---------------|---------------------------|
| | Best Local Similarity | 41.6%: | Pred. No. 2.8e-08: | | | | | |
| | Matches | 429: | Conservative | 0: | Mismatches 602: Indels 0: Gaps 0: | | | |
| QY | 1373 | ATATTAA | CATGACAA | CGAA | CATTCA | CGCAATATGCGCTTTGGAC | TAGAAATAGAAAA | 1432 |
| Db | 1267 | AAAAAAAA | NNNNNN | AAAAAAAA | NNNNNN | AAAAAAAA | NNNNNN | 1208 |
| QY | 1433 | GGTGAAT | ATATAAT | CACTTA | TATGATC | ACACCTTA | AGCTTGTGAGGAA | CCGCTCT |
| Db | 1207 | AAAAAA | AAAAAAAA | AAAAAAAA | AAAAAAAA | NNNNNN | NNNNNN | 1146 |
| QY | 1493 | GTA | AACCACTG | ATATCA | TTCAGCA | AAAAAA | CTTAATG | AGACTACCAATTTG |
| Db | 1147 | AAAAAA | AAAAAAAA | AAAAAAAA | AAAAAAAA | AAAAAAAA | AAAAAAAA | 1088 |
| QY | 1553 | ATT | TACATTTAT | TACAGAA | ATATAT | CACTAAT | TAAAGCAAT | CAACATATTATCT |
| Db | 1087 | AAAAAA | AAAAAAAA | AAAAAAAA | AAAAAAAA | NNNNNN | NNNNNN | 1028 |
| QY | 1613 | TAG | CTATCAT | CTCTTG | ATATTTCA | ACACATG | ATAAATAC | GGCTCAAAATATAGAA |
| Db | 1027 | AAAAAA | AAAAAAAA | AAAAAAAA | AAAAAAAA | AAAAAAAA | AAAAAAAA | 968 |
| QY | 1673 | TCA | ACGTTAC | CCATAT | TATTA | AAAGCAG | CAATAC | CCATCGTTACATCCAACTAG |
| Db | 967 | AAAAAA | AAAAAAAA | AAAAAAAA | AAAAAAAA | AAAAAAAA | AAAAAAAA | 908 |
| QY | 1733 | TAT | TACATAT | TTCATAC | CAACTTA | AGTTA | CTTAATTTCT | CGTCAATATATATA |
| Db | 907 | AAAAAA | AAAAAAAA | AAAAAAAA | AAAAAAAA | AAAAAAAA | AAAAAAAA | 848 |
| QY | 1793 | AAC | AAATTCCT | ATCTCT | CAATCT | CAATTA | CCAAATGTTAT | CAGTTATTAATAGTCAA |
| Db | 847 | AAAAAA | AAAAAAAA | AAAAAAAA | AAAAAAAA | AAAAAAAA | AAAAAAAA | 788 |
| QY | 1853 | TTA | ACTTTTACT | ATTTAT | TATTAAT | TATATG | TGCACTTTTCA | CATCATGATAT |
| Db | 787 | AAAAAA | NNNNNN | NNNNNN | NNNNNN | NNNNNN | NNNNNN | 728 |
| QY | 1913 | ATA | TGACAT | TAATTT | TACATA | AAATTA | CTAAT | TAATTAATTT |
| Db | 727 | AAAAAA | AAAAAAAA | AAAAAAAA | AAAAAAAA | AAAAAAAA | AAAAAAAA | 668 |
| QY | 1973 | ACT | AAAA | CACGCT | ATTTCT | ACGTAT | TTTCAATG | ATGCTCTGCACTAATTAATTTATC |
| Db | 667 | AAAAAA | NNNNNN | NNNNNN | NNNNNN | NNNNNN | NNNNNN | 608 |
| QY | 2033 | ATC | CTTGCAT | ATTTCT | CAACAT | TAATAT | GATAA | CTGTTTTTTTATGCCATTAAGTTCT |
| Db | 607 | AAAAAA | AAAAAAAA | AAAAAAAA | AAAAAAAA | AAAAAAAA | AAAAAAAA | 548 |
| QY | 2093 | TAA | TAAATTTG | CTATAT | TACCA | AAAAAG | TTAAAA | TACATTCCTGCACAAT |
| Db | 547 | AAAAAA | AAAAAAAA | AAAAAAAA | AAAAAAAA | NNNNNN | NNNNNN | 488 |
| QY | 2153 | AGA | TACATAT | CAGCTCA | ATTAAT | TAAACA | AAATTTCT | TAATATTAACA |
| Db | 487 | AAAAAA | AAAAAAAA | AAAAAAAA | AAAAAAAA | NNNNNN | NNNNNN | 428 |
| QY | 2213 | TAG | AAACCA | TATGAT | TCTACT | AGTA | AGATAT | CCCTACGACATGATCAATATATA |
| Db | 427 | AAAAAA | NNNNNN | NNNNNN | NNNNNN | NNNNNN | NNNNNN | 368 |

COMMENT

Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: segre@genoscope.cns.fr; Web: www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 7009.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0CAP004DD02QPLcluster=7009.f. Contact :
Feng Liang Email: fliang@life.com URL :
http://fulllength.invitrogen.com/ invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0CAP004DD02QPL.

FEATURES

source

1.1061
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0CAP004YH04"
/issue_type="THYMUS"
/clone_lib="Homo sapiens THYMUS"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN

Query Match

Best Local Similarity 42.4%; Score 115.8; DB 13; Length 1061;
Best Local Similarity 42.4%; Pred. No. 4.5e-08;
Matches 297; Conservative 56; Mismatches 343; Indels 4; Gaps 1;

1501 ACTGATATCATTAGCAAAAAAATTATGATAGACACACCAATTGCTATGCTATTAACA 1560
339 ATTATATTTTAAATATATTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTT 398
1561 TTTATATGAGATATATATATATATATATATATATATATATATATATATATATAT 1620
399 TTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 458
1621 CATCCTTTGATA---ATTCAACACATGATATATATATATATATATATATATATAT 1676
459 TTTATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 518
1677 ACCTTCACCATATATATATATATATATATATATATATATATATATATATATATATAT 1736
519 AATAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 578
1737 TAAATATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1796
579 AAAAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 638
1797 AATTCCTATCTCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTAT 1856
639 AAAAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 698
1857 ACTTTTACTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 1916
699 AAAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 758
1917 TGAACATATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1976
759 TATATATATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 818
1977 AAGACATATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 2036
819 AAAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 878
2037 TTGACATATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 2096
879 CAAATATATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 938
2097 AAAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 2156

Db

939 YTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 998
2157 ACAATACAGCTCAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 2196
999 ACAATATCTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1038

Db

RESULT 40

CG744200/c

LOCUS

DEFINITION

P036-3-H01.2a Ppa EcoRI BAC Library Pristionchus pacificus genomic,

genomic survey sequence.

ACCESSION

CG744200

VERSION

CG744200.1

KEYWORDS

SOURCE

ORGANISM

REFERENCE

1 (bases 1 to 1305)

AUTHORS

Strinivasan,J., Sinz,W., Jesse,T., Wiggers-Pereboite,L., Jansen,K.,

Bautjer,J., van der Meulen,W. and Sommer,R.U.

An integrated physical and genetic map of the nematode Pristionchus

pacificus

Mol. Genet. Genomics 269 (5), 715-722 (2003)

22835951

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601488
Email: ralf.sommer@tuebingen.mpg.de
Class: BAC ends.

FEATURES

source

1.1305
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Ppa EcoRI BAC Library"
/note="The library was generated by a partial digest of
the genomic DNA with EcoRI and cloning into the BAC
vector."

ORIGIN

Query Match

Best Local Similarity 42.6%; Score 115.6; DB 29; Length 1305;
Best Local Similarity 42.6%; Pred. No. 4.5e-08;
Matches 379; Conservative 0; Mismatches 511; Indels 0; Gaps 0;

1420 AGAATATGAGAAAGTAGAAATATATATATATATATATATATATATATATATATATATAT 1479
1291 AAAAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1232
1480 AGAAGACAGTACTGTAATCCACTGATCATTTAGCAAAAAAATTATATGTAAGTACTC 1539
1231 AAAAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1172
1540 AATTGATATCTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 1599
1171 AAAAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1112
1600 AAGCATATATCTTATCTATCATCTTTGATATTTTCAACATGATTAATTAACGCT 1659
1111 AAAAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1052
1660 CAAAAATATGATATCAAGCTTCACACATATATATATATATATATATATATATATATATAT 1719
1051 AAAAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 992
1720 ACATCCAGTATGATATATATATATATATATATATATATATATATATATATATATATAT 1779

